Molecular markers from the mitochondrial genome of arbuscular mycorrhizal fungi (Glomeromycota): evolutionary dynamics and application

Inauguraldissertation

zur
Erlangung der Würde eines Doktors der Philosophie
vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel

von
Odile Thiéry
aus Semur-en-Auxois, Frankreich
Basel, 2010

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät
auf Antrag von

Prof. Dr. Andres Wiemken, Prof. Dr. Dirk Redecker, Prof. Dr. Thomas Boller

Basel, den 21/09/2010

Prof. Dr. Martin Spiess, Dekan

Table of contents

Summary

The arbuscular mycorrhizal fungi (AMF) are obligate biotrophic symbionts with a key role in ecosystem functioning by their main contribution to plant mineral nutrient uptake. A fragment of the mitochondrial large subunit gene of ribosomal RNA (mtLSU) has been established as a molecular marker for population studies in *Glomus intraradices*, based on its homogeneity within isolate though genetic variation among isolates.

The respective region of the mtLSU was analyzed in five species of *Glomus* (*G. mosseae*, *G. geosporum*, *G. caledonium*, *G. clarum*, *G. coronatum*) from the same major clade (*Glomus* group A), *Glomus* sp. ISCB 34 from the related *Glomus* group B and two species of *Scutellospora*. The nucleotide polymorphism was very low among related morphospecies and differences were mainly caused by the exon/intron structure, limiting the resolution to a species level. Phylogenetic analyses suggested vertical inheritance from common ancestors for some introns as well as horizontal transfer for others.

Four mitochondrial intergenic spacers (rns/nad5, atp6/nad2, nad3/nad6, and nad6/cox3) were assessed as potential intraspecific markers. Combined, they offer a resolution even higher than the mtLSU, with the highest polymorphism found in the rns/nad5 spacer, most likely caused by processes of evolutionary dynamics involving homing endonucleases.

A mtLSU PCR approach was applied to study the population structure of *G. intraradices* at geothermal sites in Europe and North America, revealing considerable newly-discovered mtLSU sequence types as well as previously-reported types from arable fields. The molecular survey showed strong biogeographical structure in the occurrence of *G. intraradices*.

Overall, this PhD thesis project showed that the genetic characteristics of the AMF mitochondrion are promising for development of molecular genetic markers and evolutionary sequence analyses, and will therefore be relevant for further studies of genetics, ecology and evolution of AMF.

Chapter 1: General introduction

I. Mycorrhizal symbioses

The mycorrhizal symbiosis (from Greek, mykes - fungus and rhiza - root) is an association with reciprocal benefits between fungi from different taxa (Zygomycota, Glomeromycota, Ascomycota and Basidiomycota) and roots of numerous plant species from a wide spectrum (Anthocerophyta, Marchantiophyta, Bryophyta, Pteridophyta and Spermatophytes). Depending upon the type of colonization, the mycorrhizae are divided up into three categories: ectomycorrhiza, ectendomycorrhiza and endomycorrhiza (Fig. 1; Smith & Read, 1997). In the case of the ectomycorrhiza, the fungus often forms a sheat or a mantle that surrounds the root surface. The hyphae penetrate inwards between the root epidermal and cortical cells, thus forming a network called the Hartig net. In the case of the ectendomycorrhiza, the mantle may be less important or even absent. In general, the Hartig net is well-developed but the hyphae penetrate into the plant cells. In the case of the endomycorrhiza, the mantle is absent and the fungal hyphae penetrate into the plant cells. Each category of mycorrhiza includes different types which are classified according to their morphology, their physiological characteristics and the nature of the plants and fungi involved (Table 1, Fig. 1).

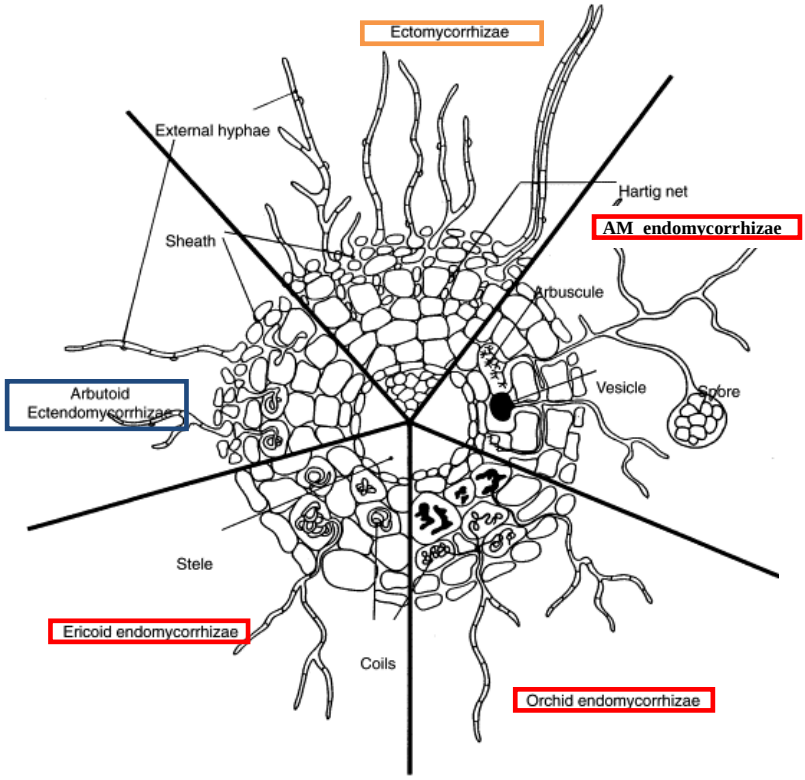


Fig. 1: Growth patterns of mycorrhiza types. From Selosse & Le Tacon (1998).

Among the types of endomycorrhiza, the arbuscular mycorrhizal (AM) symbiosis is the most common and the one that is studied in this thesis. It is an association between plants belonging to liverworts and hornworts, Pteridophytes, Gymnosperms as well as Angiosperms and fungi of the phylum Glomeromycota (Schüßler et al., 2001b). Only few plant families such as the Brassicaceae, the Chenopodiaceae, the Caryophyllaceae, and the Juncaceae do not associate with the Glomeromycota even though exceptions were found even in these families. According to fossils that have been found, Glomeromycota-like fungi date back to approximately 460 million years (Redecker et al., 2000). It is assumed that 80% of terrestrial vascular plant species form AM (Smith & Read, 1997) and it seems that this symbiosis may have played a role in the migration of plants from the aquatic to the terrestrial environment. The term arbuscular mycorrhizal symbiosis is derived from distinctive structures, the arbuscules, observed in the plants forming this type of symbiosis. The arbuscules are formed by the dichotomous branching of the

hyphae and create symbiotic interfaces in the plant cortical cells. They thus have a key role in the exchange of nutrients (Yoshida & Parniske, 2005), mainly phosphate and nitrogen to the plant and carbohydrates to the fungus (chapter1, section II). The cost of carbon allocation could be detrimental for the plants and could shift mutualism towards parasitism. Not only fungal partners but also plant hosts could be parasitic. Bidartondo et al. (2002) provided an example of non-photosynthetic plants which parasite their neighbouring plants by taking up their carbohydrates via a common fungal network.

Table 1: Characteristics of important mycorrhizal types (modified after Smith & Read, 2008). Fungal taxa are abbreviated from Glomeromycota, Zygomycota, Ascomycota and Basidiomycota; plant taxa from Bryophyta, Pteridophyta, Gymnospermae and Angiospermae. The structural characters given relate to the mature state, not the developing or senescent states. Entries in brackets indicate rare conditions. *, all orchids are achlorophyllous in the early seedling stages. Most orchid species are green as adults.

	Kinds of mycorrhizas						
	Arbuscular	Ecto-	Ectendo-	Arbutoid	Monotropoid	Ericoid	Orchid
Fungal taxa	Glomero	Basidio, Asco, (Zygo)	Basidio, Asco	Basidio	Basidio	Asco	Basidio
Plant taxa	Bryo, Pterido, Gymno, Angio	Gymno, Angio	Gymno, Angio	Ericales	Monotropa-ceae	Ericales, Bryo	Orchidaceae
Fungi septate	-	+	+	+	+	+	+
aseptate	+	-	-	-	-	-	-
Intracellular colonization	+	-	+	+	+	+	+
Fungal mantle	-	+	+ or -	+ or -	+	-	-
Hartig net	-	+	+	+	+	-	-
Achlorophylly	- (+)	-	-	-	+	-	+*

II. Development of the arbuscular mycorrhizal (AM) symbiosis

Arbuscular mycorrhizal fungi (AMF) are obligate symbionts relying on a living photoautotrophic partner to complete their life cycle. They are thought to be asexual since no sexual reproductive structures have been observed. The Glomeromycota form relatively large spores, from 22 to 1050 µm in diameter (Schenck & Perez, 1990). These spores typically contain hundreds to thousands of nuclei (Bécard & Pfeffer, 1993), large amount of lipids and carbohydrates as well as organelles, including hundreds of mitochondria (Lang & Hijri, 2009). Spores may be formed singly, in clusters or in so-

called sporocarps (Gerdemann & Trappe, 1974). Together with infected root fragments and hyphae, spores are responsible for propagation and dispersal.

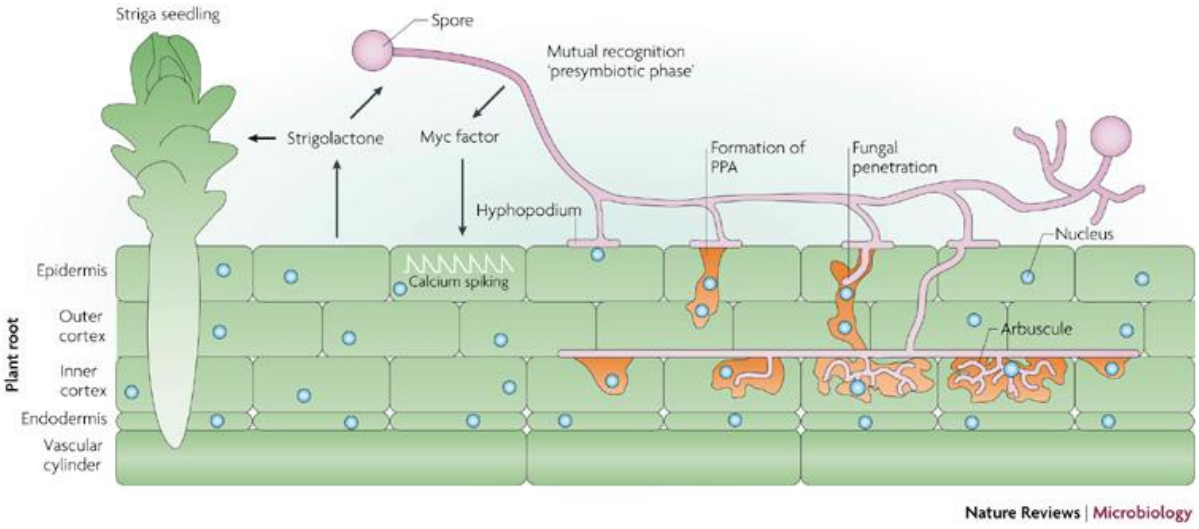


Fig. 2: Steps in arbuscular mycorrhiza (AM) development. Colonization process is depicted from left to right. Plants produce strigolactones that induce not only physiological activity of the fungus but also seed germination in parasitic plants, such as *Striga*. Fungi produce mycorrhiza "myc factors" that induce calcium spiking in root cells in order to activate plant symbiosis-related genes. AMF form hyphopodia. In response, plant cells produce a prepenetration apparatus (PPA). Then, a fungal hypha grows through root cells towards the cortex to form arbuscules. Spores are produced. From Parniske (2004).

The AM symbiosis shows several stages (Fig. 2): 1) the plant–AMF interaction is initiated by mutual signals: roots release exudates (e.g. strigolactones; Akiyama & Hayashi, 2006) that induce the resting AMF spores to germinate and stimulate hyphal growth and branching, whereas AMF produce signal molecules termed "myc factors" that induce the molecular and cellular responses of the host necessary for the establishment of the symbiosis, 2) the coenocytic AMF hyphae enter into contact with the roots where they differentiate into appressoria (hyphopodia); 3) plant cells produce a prepenetration apparatus (PPA; Genre et al., 2005), a 'transcellular tunnel' through which the hyphae penetrate into the roots and continue their growth (Parniske, 2004). Two different morphological growth patterns (Dickson et al., 2007) of AMF within the roots have been reported depending on the influence of both plant and fungal partners: in the *Arum* type,

hyphae quickly colonize the intercellular space of the root cortex and continue their growth until they differentiate into arbuscules in cortical cells; in the *Paris* type, the hyphae grow only intracellularly and form hyphal and arbuscular coils. The arbuscules (chapter 1, section I) are the sites for nutrient exchange, mainly phosphate and nitrogen are transported to the plant and carbohydrates from the photosynthesis from the plant to the fungus. Besides arbuscules, vesicles can be formed inside or between the host cells: they are filled with lipids and numerous nuclei and thought to function as energy storage organs for the fungus. Vesicles are however not formed by the Gigasporaceae; 4) simultaneously to intraradical colonisation, the fungus develops a network of hyphae in the soil: the extraradical mycelium. The exploration zone of the soil by the plants is considerably increased due to the extraradical hyphae of the AMF. In most species, sporulation takes place outside of the plant root at the tip of fungal hyphae.

III. Ecological significance of the AM symbiosis

"The study of plants without their mycorrhizas is the study of artefacts. The majority of plants, strictly speaking, do not have roots; they have mycorrhizas." This statement from the International Bank for the Glomeromycota (IBG) committee reflects the ecological importance of the AM symbiosis. AMF are present in complex ecosystems, such as tropical rainforests and temperate grasslands, and less diverse in highly disturbed, nutrient-rich, arid or very wet habitats (Smith & Read, 1997). AMF have a considerable effect on ecosystem functioning. They enhance the supply of phosphorus, the limiting nutrient in many ecosystems, to plants and influence biogeochemical cycles (Scheublin et al., 2004). They are thought to influence microbial diversity (Johnson et al., 2004), plant diversity and plant productivity (van der Heijden et al., 1998), and seedling development by their regulation of nutrient, carbohydrate and water fluxes. They play an important role in ecosystem development and succession as they form common mycelial networks: the same fungus can colonize several plants and a plant can be a host for up to 20 AMF sequence types (Fitter et al., 2005). Host specificity was thought to be very low (Smith & Read, 1997) but some studies revealed host preference especially in mycoheterotrophic mycorrhizas (e.g. Bidartondo et al., 2002) but also in autotrophic plants (e.g. Sýkorová et al., 2007b). Generally, the mycelial network of AMF increases not only plant biomass but

also resistance to biotic (e.g. nematodes, pathogens) and abiotic stresses (e.g. heavy metals) and consequently improves plant fitness. This leads to higher crop yields especially in nutrient-poor soils (Klironomos et al., 2000) and could be used as an alternative to a high input of chemical fertilizers and pesticides in agricultural practice. AMF are effective in stabilization of soil aggregates (Schreiner & Bethlenfalvay, 1995) through glomalin production, erosion control and water management (Piotrowski et al., 2004).

IV. Classification of the AM fungi (AMF)

So far, about 200 species of AMF have been described (http://www.lrz-muenchen.de/~schuessler/amphylo/amphylo_species.html). AMF are classified into four orders encompassing ten families and fifteen genera (Table 2). *Glomus* is the largest genus with approximately half of the described species. This non-monophyletic genus has been subdivided into three subgroups *Glomus* groups A, B and C (Schwarzott et al., 2001). *Glomus* group C is in the process of being reclassified into the genus *Diversispora*, order Diversisporales.

Table 2: Classification of AMF in orders, families and genera. Modified from http://www.lrz-muenchen.de/~schuessler/amphylo/.

orders (4)	**families** (10)	**genera** (15)	**litterature**
Glomerales	*Glomeraceae*	*Glomus*	Tul. & C. Tul. (1845)
Diversisporales	*Gigasporaceae*	*Gigaspora*, *Scutellospora*, *Racocetra*	Gerd. & Trappe (1974), C. Walker & F.E. Sanders (1986), Oehl, F.A. Souza & Sieverd. (2008)
	Acaulosporaceae	*Acaulospora* & *Kuklospora*	Trappe & Gerd. (1974), Oehl & Sieverd. (2006)
	Entrophosporaceae	*Entrophospora* (unclear phylogenetic affiliation)	R.N. Ames & R.W. Schneid. (1979)
	Pacisporaceae	*Pacispora*	Oehl & Sieverding (2004)
	Diversisporaceae	*Diversispora* & *Otospora* (unclear phylogenetic affiliation)	C. Walker & A. Schuessler (2004), J. Palenzuela, N. Ferrol & Oehl (2008)
Paraglomerales	*Paraglomeraceae*	*Paraglomus*	J.B. Morton & D. Redecker (2001)
Archaeosporales	*Geosiphonaceae*	*Geosiphon*	F. v. Wettstein (1915)
	Ambisporaceae	*Ambispora*	C. Walker, Vestberg & Schuessler (2007)
	Archaeosporaceae	*Archaeospora* & *Intraspora* (very closely related genera)	J.B. Morton & D. Redecker (2001), Oehl & Sieverd. (2006)

Traditionally, AMF species have been distinguished morphologically by features of the spore wall. Genera and families were based on the mode of spore formation (the way the spore is formed on the hypha; Fig. 3) and species on the layered structure of the spore walls (Morton, 1988). Wall components were grouped hierarchically taking in consideration spore development (Morton et al., 1997; Stürmer & Morton, 1999). Descriptions found on the INVAM website (http://invam.caf.wvu.edu) provide a reference for species identification. Nevertheless, identification of species using morphological characters can be challenging if spores are parasitized, degraded, at different life stages (young and old spores may vary morphologically in color, for example) or the species are dimorphic. Moreover, spore surveys do not reveal species which are not sporulating at the investigation time and do not necessarily reflect symbiotically active (i.e. root-colonizing) fungi. Consequently, alternative molecular approaches were developed. Some of these used biochemical characteristics such as fatty acids methyl esters (Graham et al., 1995; Bentivenga & Morton, 1996), others used nucleic acid-based techniques (Lloyd MacGilp et al., 1996; Helgason et al., 1998a). DNA-based phylogenetic methods have also been successfully used to elucidate the position of AMF within the Fungi and a new monophyletic fungal phylum, the Glomeromycota, was established based on phylogenetic analyses of the rDNA small subunit (SSU; Fig. 3; chapter1, section VIII; Schüßler et al., 2001b). This phylum also comprises *Geosiphon pyriformis*, the only known member forming a symbiosis with the cyanobacterium *Nostoc punctiforme*.

In this common phylogeny, the phylum Glomeromycota was defined as a sister group of the Dikarya, i.e. Ascomycota and Basidiomycota, which are characterized by regularly septate hyphae and a dikaryotic phase in their life cycle. A six-gene phylogeny based on three rRNA genes (the large subunit [LSU], the Internal Transcribed Spacers [ITS] and the small subunit [SSU]), the RNA polymerase II subunit 1 and 2 (RPB1 and 2), elongation factor 1 alpha (EF1α) also supported a clade uniting the Dikarya and the Glomeromycota (James et al., 2006). Tehler et al. (2003) even proposed a novel phylum "Symbiomycota" comprising Glomeromycetes, Ascomycetes and Basidiomycetes. However, phylogenies based on protein-coding genes (e.g. RPB1; Redecker & Raab, 2006) and the mitochondrial genome sequence of *Glomus intraradices* isolate FACE494

(Lee & Young, 2009) defined the Glomeromycota as one of the lineages of the paraphyletic Zygomycota (Lee & Young, 2009; Liu et al., 2009).

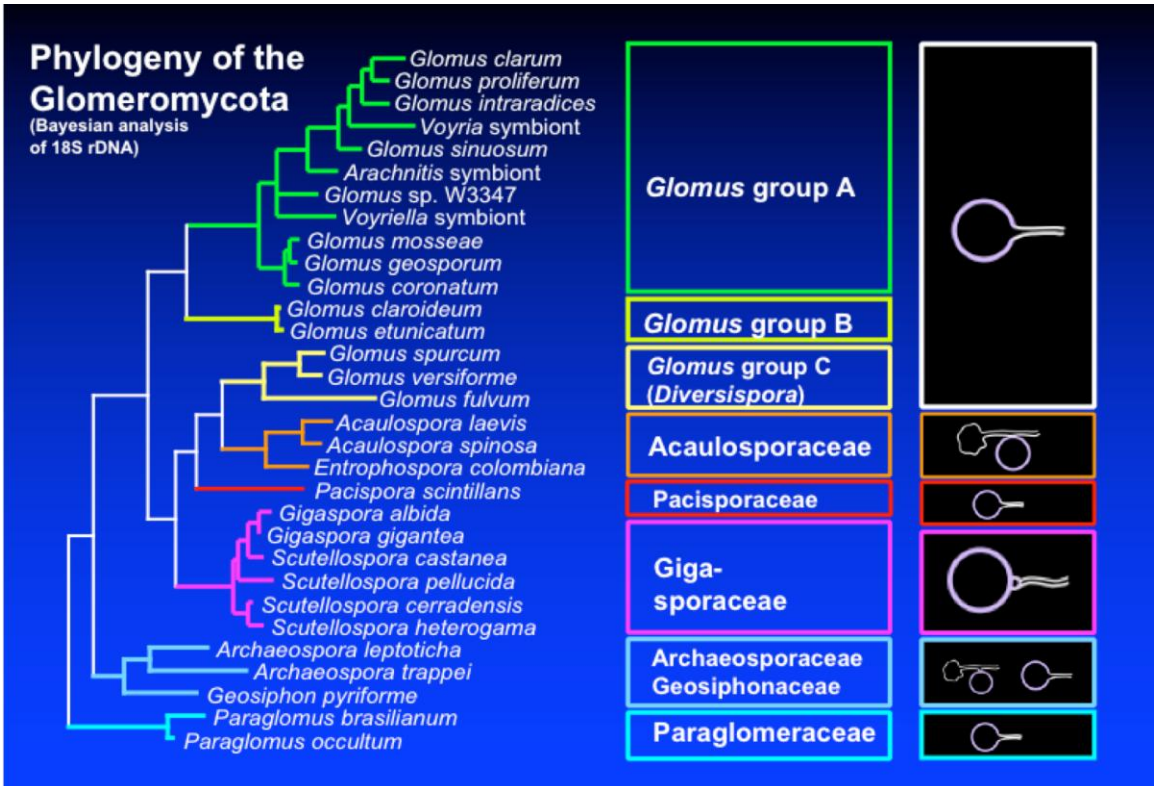


Fig. 3: Phylogeny of the main Glomeromycotan groups based on Bayesian analysis of the 18 S rDNA. Mode of spore formation is depicted in the boxes. © Dirk Redecker.

V. Complexity of genetics in AMF

1. Nuclear genome

In the nuclear genome, unusual polymorphism of ribosomal DNA and protein-coding genes was observed within single spores of AMF (Sanders et al., 1995). Two major hypotheses have been proposed to explain this genetic variation (Pawlowska & Taylor, 2004; Hijri & Sanders, 2005):

i) Heterokaryosis: the variants of a locus may be distributed among different nuclei (Fig. 4a).

ii) Homokaryosis: all variants exist in each nucleus due to polyploidy (Fig. 4b) or due to haploidy (Fig. 4c).

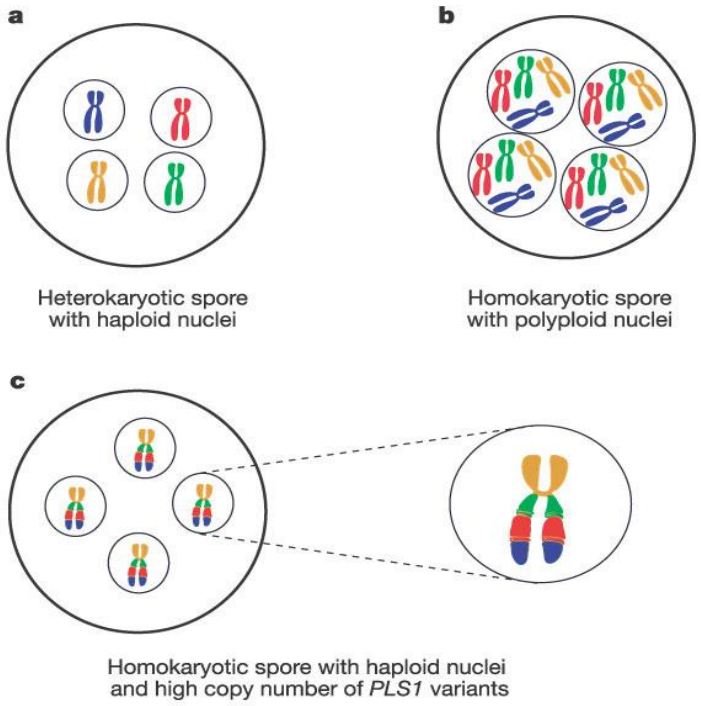


Fig. 4a–c: Hypotheses explaining the DNA polymorphism observed in spores of *G. etunicatum*. PLS1 variants exist in different nuclei (a; heterokaryosis), all PLS1 variants are present in each nucleus owing to polyploidy and nuclei are genetically identical (b; homokaryosis), or all PLS1 variants are present in each nucleus as copies due to duplication events of PLS1 in a haploid genome (c; also homokaryotic). For clarity we used only four colours to represent the 13 different PLS1 variants and only four nuclei instead of 750 nuclei contained in each *G. etunicatum* spore. For simplicity, haploid nuclei are depicted with one chromosome and polyploid nuclei with four chromosomes, although actual chromosome number in this fungus is unknown. The four colours represent variants of the PLS1 region, although up to 13 variants have previously been recorded per spore. From Hijri & Sanders (2005).

Based on mathematical models (Pawlowska & Taylor, 2004) favoured the homokaryosis hypothesis involving polyploidy. They established their model using the gene encoding the catalytic subunit of DNA polymerase α (POL1) which typically occurs as a single copy in the genome; this gene was found to be falling into two distinct phylogenetic clusters of POL1-like sequences (PLS), PLS1 and PLS2. The authors demonstrated inheritance of all 13 variants of PLS1 without any loss in five daughter spores in *Glomus etunicatum*, which would be unlikely to occur in a heterokaryotic system due to bottleneck events. Hijri & Sanders (2005) estimated the total nuclear DNA

content of *G. etunicatum* to be 37.45 Mb by flow cytometry. This genome size would rule out the 13-fold polyploidy required by Pawlowska & Taylor's model, because it would imply a genome size of 2.88 Mb, smaller than that of most eukaryotes and bacteria. Instead, Hijri & Sanders (2005) determined that *G. etunicatum* contains 1.88 copies of the PLS1 per nucleus by real-time PCR and has a haploid genome by DNA-DNA fluorescent in situ hybridization. Bever & Wang (2005) argued that distribution of genetic variation could also be maintained in a heterokaryotic system by fusion of hyphae (anastomosis) as observed within isolates of *Glomus mosseae*, *Glomus intraradices* and *Glomus proliferum* (Giovannetti et al., 2004; Voets et al., 2006). Recently, Croll et al. (2009) even showed genetic exchange through anastomoses among different isolates of *G. intraradices* and Croll & Sanders (2009) demonstrated recombination in the genome of a subset of isolates. Nonetheless, gene exchanges in AMF are thought to be very rare (den Bakker et al., 2010). The whole genome of *G. intraradices* isolate DAOM197198 is currently being sequenced (Martin et al., 2008) and will help to elucidate the probably very unusual genetics of AMF.

2. Mitochondrial genome

Recently, Lee & Young (2009) have sequenced for the first time the mitochondrial genome of an AMF, *G. intraradices* isolate FACE494 (Fig. 5). In fungi, mitochondrial genomes (chondriomes) range from 18-175 kb in size (*Schizosaccharomyces pombe* and *Agaricus bitorquis* respectively), which is the intermediate size between mitochondrial genomes in animals (approximately 16kb) and plants (130-350 kb). Lee & Young (2009) determined the size of the chondriome of *G. intraradices* to be 70606 bp. Fungal chondriomes are generally circular and have a lower G + C content than the nuclear genomes, with a G + C content of 37.2% in *G. intraradices* (Lee & Young, 2009). Gene order in fungal mitochondria is not conserved among major lineages (e.g. *Neurospora crassa* and *Aspergillus nidulans*). Nevertheless, they typically encode ubiquitous gene products involved either (i) in the mitochondrial translational apparatus (tRNAs, the small subunit (*rns*) and the large subunit (*rnl*) ribosomal RNAs), or (ii) in the oxidative phosphorylation (7 subunits of the NADH dehydrogenase complex (*nad*), apocytochrome

b (*cob*), subunits 1-3 of cytochrome c oxidase (*cox*) and subunit 6, 8 and 9 of ATP synthase (*atp*)). The triplet coding corresponds to the standard genetic code except that UGA is assigned not to a stop codon but to tryptophan, as in many fungal mitochondria (i.e. genetic code 4).

Twenty-six introns were reported in the mitochondrial genes of *G. intraradices*, 23 belonging to group I introns and 3 uncharacterized so far. Group I introns are defined through four blocks of conserved sequences and additional complementary non-conserved sequences. They are self-splicing *in vivo*, but require nuclear-encoded proteins *in vitro*. This group of introns is potentially mobile, and its mobility can be promoted by homing endonucleases (HE) encoded by open reading frames (ORFs) within the introns.

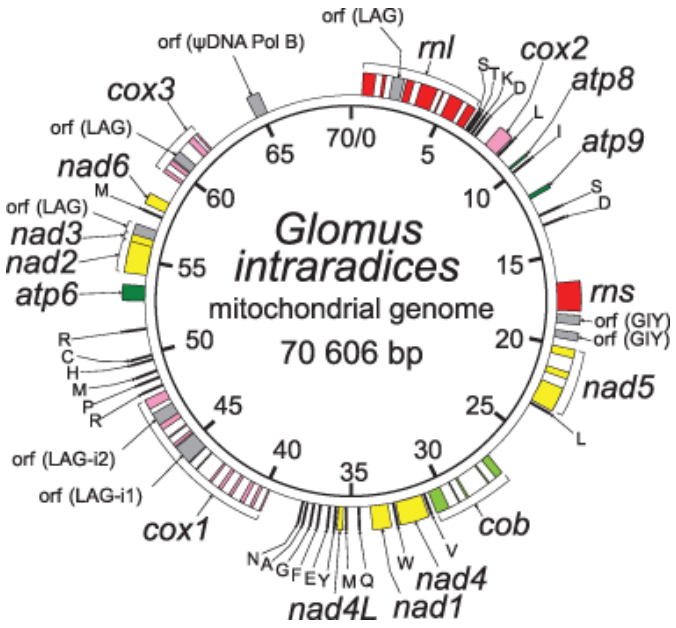


Fig. 5: The mitochondrial genome of *G. intraradices* isolate FACE494. The outer circle shows the location of genes (shaded) and introns (white); all genes are oriented clockwise. The scale is in kb. From Lee & Young (2009).

VI. Mitochondrial inheritance

Mitochondrial inheritance differs from nuclear inheritance. In most sexually-reproducing eukaryotes, mitochondrial genomes are inherited from a single parent (Sears, 1980; Birky et al., 1983) whereas nuclear genomes originate from two parents. Biparental inheritance and recombination confer advantages in terms of repairing DNA damage and purging of deleterious mutations that do not apply to mitochondrial genomes (Birky, 1995). Indeed, multiple copies of the mitochondrial genome are contained in many mitochondria and recombination of nuclear genomes to repair DNA damage is not essential for mitochondrial DNA (mtDNA).

In fungi, the mechanisms that exclude mitochondria from one parent occur at different stages (pre-zygotic, fertilization and zygotic) of the sexual cycle (Xu, 2005):

-At the pre-zygotic stage, the paternal and maternal gametes often do not contain the same number of mitochondrial genomes during gametogenesis. The male gamete possesses fewer mitochondria than the female gamete and its mtDNA is hardly detectable in progenies. This is the case in many Ascomycetes such as *Neurospora crassa*.

-At the fertilization stage, mtDNA from one gamete fails to penetrate the zygote after hyphal fusion. Such mechanisms are widespread in Basidiomycetes.

-At the zygotic stage, stochastic transmission, selective replication, or degradation of mitochondria from one parent leads to uniparental mitochondria inheritance. (i) Stochastic transmission: in zygotes, most of the mitochondria from the two parents remain in their initial position and do not completely mix. Buds arising at the end of the zygote contain in majority mtDNA from one parent whereas buds produced at the middle position typically possess mitochondria from the two parents. This process was reported in *Saccharomyces cerevisiae*. (ii) Selective replication: e.g. mating between hypersuppressive "petite" (HS) and wild type *S. cerevisiae* strains leads to uniparental mtDNA due to the high density of replication origins in HS. (iii) Degradation of mitochondria: after zygote formation, mechanisms ensure the destruction of mitochondria from one parent. This process was also described in many non-fungal eukaryotes.

Homogeneity of mitochondrial markers in AMF was shown by Raab et al. (2005; chapter 1, section VIII) and confirmed in following studies (Börstler et al., 2008; Lee & Young, 2009; Börstler et al., 2010). However, this is surprising as Giovannetti et al.

(1999) and Croll et al. (2009) reported reciprocal or unilateral genetic exchange through mycelial anastomoses. Possible processes to restore mitochondrial homogeneity could be (i) segregation mechanisms of haplotypes into different parts of the fungal mycelium (Shibata & Ling, 2007), (ii) genetic bottleneck: coexistence of heterologous mitochondrial haplotypes may be short living with a single haplotype becoming quickly dominant (Marinoni et al., 1999). To clarify this issue, large-scale genetic analyses of multiple fungal isolates should be performed to detect possible recombinations in mitochondria or occasional heteroplasmy.

VII. Homing endonucleases

1. "Homing" process

In eukaryotic cells, homing endonuclease genes (HEGs) are considered "selfish" mobile elements because their gene products typically do not participate in the survival or reproduction of the organism but rather promote their self-propagation. Usually, HEGs are encoded in self-splicing introns (chapter 1, section V.2.) and do not disrupt the function of the host gene. Their spread through populations is achieved by a process termed "homing" (Fig. 6). Homing refers to the faithful insertion in a site-specific gene, the home. The HEGs encode a HE that recognizes and binds a 15-30 bp long DNA sequence occurring only once in the host genome. However, mutations of many base pairs of the recognition site do not prevent homing from occurring between closely related species as the success of HEs increases with their spread. HEs make a double-strand break in the DNA target lacking the HEG (HEG^- chromosome) and thus disrupt the recognition site. The broken chromosome is typically repaired via recombination using the homologous chromosome bearing a HEG (HEG^+ chromosome) as a template. In this way, the HEG is copied in the HEG^- chromosome, converting it into a HEG^+ chromosome (Burt & Koufopanou, 2004). "Homing" depends on the frequency with which the HEG^- and HEG^+ chromosomes encounter each other. Once the HEGs are present in each individual of a population, the selection pressure maintaining the functionality of HEs disappears. This explains the ''homing cycle" which is the spread, fixation and degeneration of HEGs within a host gene (Goddard & Burt, 1999).

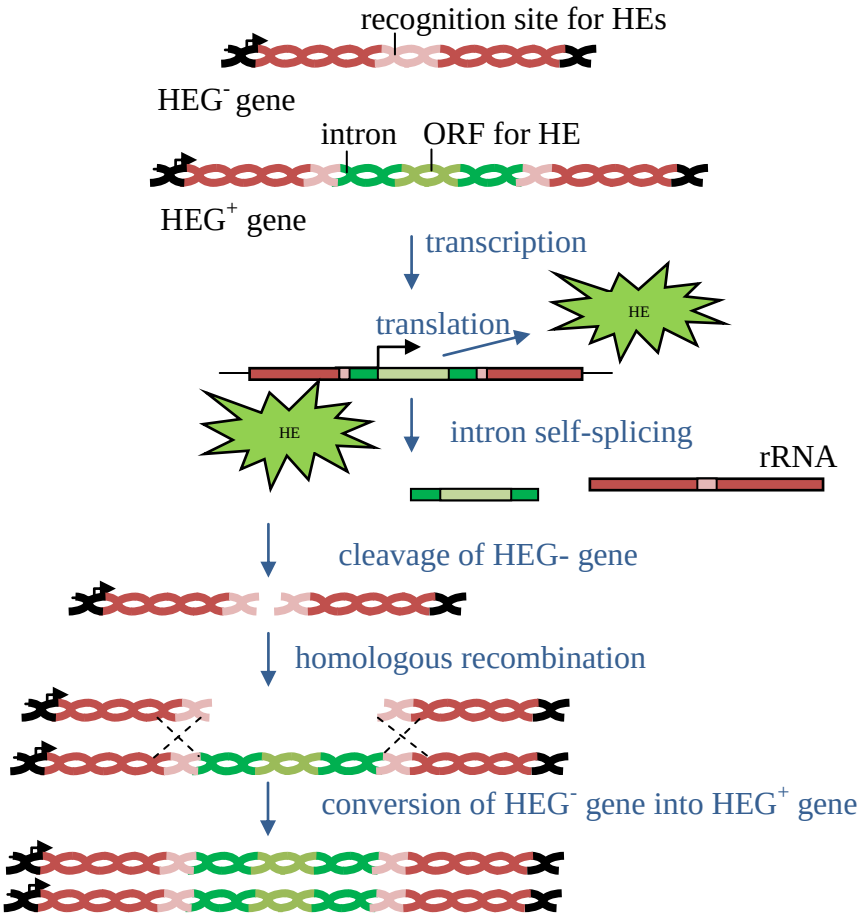


Fig. 6: Mechanisms of homing in group I introns. Based on Chevalier & Stoddard (2001).

2. Homing endonuclease families

Homing endonucleases are ubiquitous restriction enzymes present in the three domains of life -the archaea, bacteria, and eukarya- and expressed in the different compartments of the eukaryotic cell –nuclei, mitochondria, and chloroplasts. They are classified into four families based on their conserved amino-acid motifs: LAGLIDADG, His-Cys box, GIY-YIG and H-N-H (Fig. 7; Jurica & Stoddard, 1999).

a. LAGLIDADG family

This is the largest and most frequently reported family with more than 200 members up to now. The sequence motif LAGLIDADG is present in one or two motifs. The single-motif HEs act as homodimers whereas the double-motif HEs are active as monomers in such a way that the motifs create a saddle interacting with the minor groove of DNA. These motifs are directly involved in the catalysis, i.e. the DNA cleavage.

b. His-Cys box family

Two histidines and three cysteins are conserved in a region of 30 amino acids, involved in the coordination of metallic cations required for the DNA cleavage.

c. GIY-YIG family

The sequence motif GIY-YIG is present in one motif and acts as monomer. The enzyme consists of the catalytic N-terminal region comprising the conserved motif and the DNA-binding C-terminal region.

d. H-N-H family

The conserved sequence consists of 30 amino acids with two pairs of histidines and one asparagine forming a zinc finger domain.

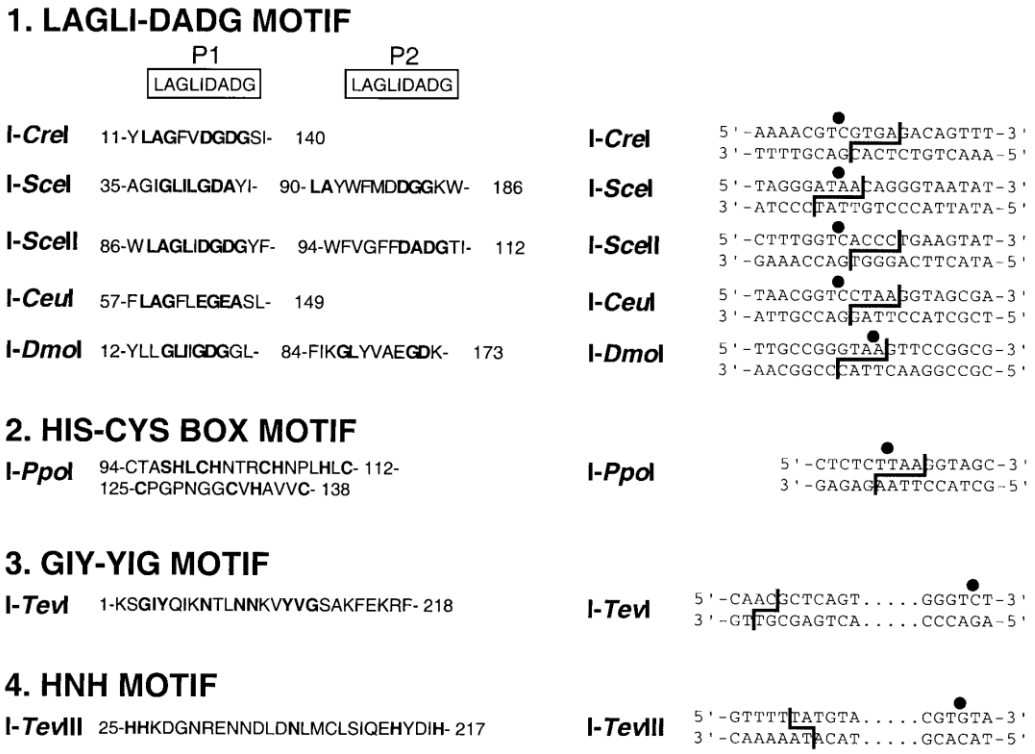


Fig. 7: Representative homing endonucleases encoded by ORFs found in group I introns. Four families have been characterized. Shown on the left are examples of the conserved namesake motifs for these proteins, and on the right the length, sequence, cleavage patterns and intron insertion site (closed circles) for their DNA target sites. From Jurica & Stoddard (1999).

VIII. Molecular markers available in AMF

Molecular tools have been used to study the genetic diversity among or within AMF species (chapter 1, section IX). Sequences of the nuclear-encoded rRNA genes are widely used in taxonomy and molecular phylogeny as they can be amplified from small quantities of DNA, due to the high copy number of rRNA genes in nucleus. Protein-coding genes are less frequently employed as molecular markers (Xu, 2005) for the following reasons: (i) the triplet codon structure makes it difficult to design robust primers across a broad range of taxa, (ii) heterozygous loci (e.g. from dikaryotic organisms such as Basidiomycota) require cloning, (iii) paralogy and gene families complicate the use of protein-encoding genes as markers.

The nuclear rRNA genes encode for two subunits (Fig. 8): the small subunit (18S [SSU]), and the large subunit containing three rRNA species (5S, 5.8S and 28S [LSU]) separated from each other by the Internal Transcribed Spacers (ITS). The 5S subunit is not normally located within this ribosomal tandem repeat. The SSU, the LSU and the ITS evolve at different rates: ITS are variable regions which change more frequently than the SSU and the LSU. This high degree of polymorphism can be explained by the relatively low evolutionary pressure on these ITS sequences that are not included in the structure of the mature ribosomes, but spliced during the rRNA maturation (Calonje et al., 2009). Nevertheless, ITS secondary structure comprises necessary sites for processing the rRNA transcripts, denoting thus a certain degree of conservation. For all these reasons, the ITS region is probably the most commonly sequenced DNA region in fungi for analyses of closely related taxa (Redecker, 2000). Nonetheless, its heterogeneity in AMF causes problems to distinguish closely related taxa or strains (Antoniolli et al., 2000; Jansa et al., 2002b). The SSU gene (Helgason et al., 1998b; Helgason et al., 1999) and the LSU gene (van Tuinen et al., 1998; Kjøller & Rosendahl, 2000; Wu et al., 2007) are usually analyzed to understand more distant relationships on the species/genus/order level as they are more conserved. Sets of primers were developed to target the SSU-ITS region (Redecker, 2000) or even a large portion of the ribosomal repeat (Fig. 8; Krüger et al., 2009).

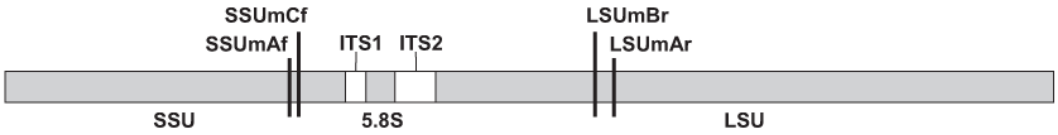


Fig. 8: Small subunit (SSU) rDNA, Internal Transcribed Spacer (ITS) region and large subunit (LSU) rDNA (5465 bp) of *G. intraradices* DAOM197198 (AFTOL-ID48, other culture/voucher identifiers: MUCL43194, DAOM181602; accession numbers: AY635831, AY997052, DQ273790) showing the binding sites of the newly designed forward and reverse primer mixtures. From Krüger et al. (2009).

The problems of apparent heterogeneity of nuclear-encoded genes within the organism and possible heterokaryotism could be circumvented by using an independent genetic system within the fungal organism, the mitochondria. In 2005, around 25 complete or nearly complete mitochondrial genome sequences were available from fungi, mainly from human or plant pathogens and other model organisms (Xu, 2005). In this

context, Raab et al. (2005) analyzed mitochondrial ribosomal RNA large subunit gene (mtLSU) sequences of *G. intraradices* and *G. proliferum* (Fig. 9) which were demonstrated to lack polymorphism within the organism (Raab et al., 2005; Börstler et al., 2008). This homogeneity of mitochondrial sequences was confirmed for the mitochondrial gene cox1 in a PhD study (Borriello, 2010) and even for the whole mitochondrial genome of *G. intraradices* after sequencing 24 spores of the isolate FACE494 (Lee & Young, 2009). The exon phylogeny of a region of the mtLSU showed superior resolution among subclades of *G. intraradices* compared to nuclear-encoded rDNA ITS (Börstler et al., 2008). Particularly, the mtLSU introns were shown to be highly sensitive molecular markers to genotype different isolates of *G. intraradices* (*sensu lato*) and it was used to differentiate mtLSU haplotypes directly from colonized roots (Börstler et al., 2008; appendix 1 and 2; chapter 1, section X), which is a promising approach to better understand the diversity and dynamics of field communities and populations of AMF. Croll et al. (2008b) also used introns of the mtLSU as a marker combined with 10 Simple Sequence Repeat (SSR) loci and introns of a nuclear gene.

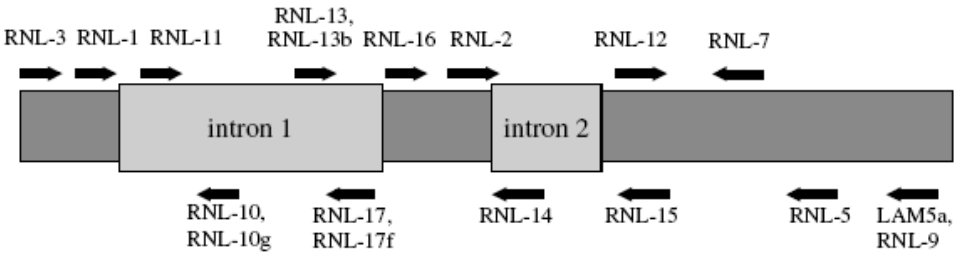


Fig. 9: Arrows show the location and orientation of the primers in the 2900 bp fragment of the mitochondrial large subunit from *G. intraradices*. Approximately to scale. From Raab et al. (2005).

IX. Fungal species concepts

Agapow et al. (2004) defined a species as "a lineage that occupies an adaptative zone minimally different from that of any other lineage in its range and that evolves separately from all lineages outside of its range". This definition is based on ecological niches and the capacity of fungi to develop and be maintained under certain evolutionary constraints. Under this concept, it is assumed that a species is characterized by particular traits

involved in biological adaptation (Xu, 2005). Ecological and physiological differences among AMF species have been observed: ecological differences were reported in terms of nutrient acquisition and transfer to the host plants (Smith et al., 2004) and physiological differences are compiled on the INVAM website (chapter 1, section IV). As a matter of fact, AMF species are largely defined only by morphology with occasionally some support from molecular data. The diversity of AMF is probably strongly underestimated since AMF field community studies revealed many newly-discovered phylotypes (e.g. Wubet et al., 2004; Börstler et al., 2006) which could not be assigned to known morphospecies. This underestimation of diversity suggests that many fungi are still undescribed and unculturable, which is highly compatible with the fungal diversity estimate of 1.5 million species (Hawksworth, 2001). Several studies pointed out differences of benefits for the plants depending on combinations of AMF species and plant hosts (e.g. Burleigh et al., 2002; Klironomos, 2003; van der Heijden et al., 2004). The differences in functionality were not restricted to different species but also observed even within species, e.g. among different AMF isolates for *G. mosseae*, *G. caledonium*, *G. geosporum*, *G. claroideum* (Munkvold et al., 2004) and for *G. intraradices* (Koch et al., 2006; Croll et al., 2008b).

X. Population studies

A population could be defined as "a group of organisms of the same species occupying a particular space at a particular time" (Krebs, 1994) and "having the opportunity to interact with each other" (Waples & Gaggiotti, 2006). Attempts were made in order to find appropriate molecular markers to study populations (chapter 1, section VIII): Amplified Fragment Length Polymorphism (AFLP) was used to differentiate four isolates of *G. intraradices* cultured in root organ cultures (ROCs; Koch et al., 2004); rDNA PCR-DGGE (Denaturing Gradient Gel Electrophoresis) patterns allowed distinction of geographic isolates of some *Gigaspora* species (de Souza et al., 2004); SSR were developed by two teams investigating either numerous isolates (48) within one field in Switzerland (Croll et al., 2008b) or few isolates from seven sites in four countries (Mathimaran et al., 2008a); nested multiplex PCR based on multiple co-

dominant genetic markers derived from single copy genes *GmFOX2* (encodes a multifunctional protein of the peroximal β-oxidation), *GmTOR2* (encodes a protein involved in cell cycle processes) and *GmGIN1* (unknown role) and LSU rDNA were developed by Stukenbrock & Rosendahl (2005b) from single spores of *G. mosseae*, *G. caledonium*, and *G. geosporum*; a PCR-RFLP-sequencing approach based on the mtLSU was established by Börstler et al. (2008) and applied to a set of 16 *G. intraradices* isolates originating from five continents (appendix 1).

Only a few studies investigating the intraspecific population structures of AMF directly from the field have been published. Stukenbrock & Rosendahl (2005a) compared the genetic structure of spores of three *Glomus* species (*G. mosseae, G. caledonium* and *G. geosporum*) from an organically and a conventionally cultivated field using their multiplex nested PCR. These authors estimate genetic diversity to be similar, with no evidence of population subdivision between the two fields but rather a population subdivision within each field. Similarly, Rosendahl & Matzen (2008) analyzed spores of the same three *Glomus* species from a fallow and a cultivated field using the same markers and showed a subdivision of *G. mosseae* haplotypes between the two fields, not observed for *G. caledonium* and *G. geosporum*. The sequence data suggested that abundance and population structure of AMF were affected by agricultural practices. Subsequently, Rosendahl et al. (2009) analyzed 82 isolates of *G. mosseae* originating from six continents to explain its worldwide distribution using the same markers. Their results indicated no geographical structure as identical genotypes were found on different continents. The authors concluded that diversification of *G. mosseae* took place after separation of the continental drift and that the lack of population structure of the fungus was due to human activity. Börstler et al. (2010) were the first to analyze the population structure of an AMF, *G. intraradices*, directly from colonized roots, that is symbiotically active fungi (appendix 2). For this purpose, they performed the PCR-RFLP-sequencing approach based on the mtLSU (Börstler et al., 2008). They investigated *G. intraradices* diversity from two agricultural field experiments in Switzerland and two semi-natural grasslands in Switzerland and France. RFLP type composition was shown to differ on one hand between agricultural and semi-natural sites and on the other hand between the two agricultural sites. Interestingly, most haplotypes from the grassland sites fell within a

separate clade, which might represent ecotypes or even different "cryptic" species. Indeed, mtLSU markers seem adequate to differentiate ecotypes as their polymorphism was correlated with symbiotic properties and growth parameters (Koch et al., 2006; Croll et al., 2008b).

XI. Aims of my thesis

In order to better understand the significance of genetic diversity within AMF species, the objective of this thesis was to explore the potential of mitochondrial molecular markers in the Glomeromycota (i) by studying the evolutionary dynamics of mtLSU and its introns in *Glomus* and *Scutellospora* species, (ii) by assessing the polymorphism of additional loci of the mitochondrial genome as new markers and (iii) by applying the mtLSU markers in field studies of the genetic structure of *G. intraradices*.

Building on previous work of Philipp Raab on *G. intraradices* and *G. proliferum* in the laboratory (Raab, 2007) and in collaboration with my colleague Boris Börstler who focused on the diversity of *G. intraradices* in the field and the development of the mtLSU as a marker for this purpose, my task was to extend the range of mitochondrial markers to other regions of the mitochondrial genome and other fungal taxa.

In chapter 2, I present my work on evolutionary dynamics and evolutionary history of the mtLSU region and its introns. These data were expected to provide a better understanding of the stability of these molecular markers and the degree of polymorphism among AMF species. Another aim of this work was to elucidate whether the lack of polymorphism in the mtLSU within fungal isolates also holds true for other species and whether it can be used as a strain-specific marker.

The question whether other regions of the mitochondrial genome are potentially useful as molecular markers is addressed in chapter 3. Using the published genome sequence of *G. intraradices*, potentially polymorphic regions were identified and, using primers designed based on these data, sequences of intergenic regions of the mitochondrial genome were obtained from several isolates of this species, showing some polymorphism.

Chapter 4 presents results from a field study of *G. intraradices* populations in geothermal and adjacent non-geothermal sites in Iceland and Yellowstone National Park (YNP). Using the method developed and optimized by Börstler et al. (2008, 2010), this work represents the logical continuation of the MSc thesis work of Susann Appoloni (2006) who analyzed AMF communities, mainly in YNP, on the species level. In the context of the surprising ubiquity of *G. intraradices*, the question arose from the results of that previous study, whether the strains of the ubiquitous species *G. intraradices* found in the geothermal areas were specialized ecotypes or the same ones as in other habitats, and whether the populations showed a geographical structure. In this regard the chapter also continues an aspect brought up by Börstler et al. (2010).

The papers by Börstler et al. (2008, 2010) are presented in the appendixes. I contributed to them by cooperating routinely with Boris Börstler. In particular, we cooperated on the amplification and characterization of the mtLSU region of *G. clarum* isolates.

Chapter 2: Evolutionary dynamics of introns and homing endonuclease ORFs in a region of the large subunit of the mitochondrial rRNA in *Glomus* species (arbuscular mycorrhizal fungi, Glomeromycota)

Published in Molecular Phylogenetics and Evolution [2010] **55**: 599–610
Odile Thiéry, Boris Börstler, Kurt Ineichen and Dirk Redecker

I. Abstract

The large subunit of the mitochondrial ribosomal RNA genes (mtLSU) has previously been identified as a highly sensitive molecular marker for intraspecies diversity in the arbuscular mycorrhizal fungus (AMF) *Glomus intraradices*. In this study, the respective region was analyzed in five species of *Glomus* (*G. mosseae, G. geosporum, G. caledonium, G. clarum, G. coronatum*) from the same major clade (*Glomus* group A), *Glomus* sp. ISCB 34 from the related *Glomus* group B and two species of *Scutellospora*. Results show low level of genetic polymorphism between related morphospecies. Introns homologous to those found in *G. intraradices* were detected as well as new ones, some of them containing putative ORFs for homing endonucleases (HEs). Introns without ORFs for HEs seem to have been inherited strictly vertically from the ancestors of *Glomus* groups A and B while other introns indicate occasional horizontal transfer and possibly maintenance, degeneration and loss together with their associated HE ORFs. Overall, we provide first insights into the evolutionary dynamics of introns and HEs in this ecologically important group of fungi, which was previously not analyzed in this respect.

Keywords: mitochondrial large subunit RNA genes, molecular markers, homing endonucleases, group 1 introns, *Glomus* group A, Glomeromycota

II. Introduction

Arbuscular mycorrhiza is a ubiquitous symbiosis between the large majority of land plants and fungi from the phylum Glomeromycota. The fungal partners are obligate symbionts and their genetics has been the subject of a lively debate, with some authors claiming genetic heterogeneity between the nuclei in the coenocytic mycelium (Kuhn et al., 2001; Koch et al., 2004; Hijri & Sanders, 2005), which was disputed by others (Pawlowska & Taylor, 2004).

In contrast to rDNA sequences and other genes from the nuclear genome (Sanders et al., 1995; Lloyd MacGilp et al., 1996), mitochondrial large subunit rDNA sequences (mtLSU) of the arbuscular mycorrhizal fungi (AMF) *Glomus intraradices* and *Glomus proliferum* were demonstrated to lack polymorphism within the same strain (Raab et al., 2005; Börstler et al., 2008). The recently sequenced mitochondrial genome of *G. intraradices* confirmed this homogeneity for the whole genome (Lee & Young, 2009).

The mtLSU and in particular its introns were shown to be highly sensitive molecular markers to genotype different isolates of *G. intraradices* (*sensu lato*) and it was used to differentiate mtLSU haplotypes directly from colonized field-collected roots (Börstler et al., 2008), which is a promising approach to obtain a better understanding of the diversity and dynamics of field communities and populations of AMF. The exon phylogeny of a region of the mtLSU showed superior resolution among subclades of *G. intraradices* compared to nuclear-encoded rDNA internal transcribed spacers (Börstler et al., 2008).

Mitochondrial DNA has a long history as a molecular marker that extends into the era before PCR facilitated the access to its sequences from a broad range of organisms (e.g. Bruns et al., 1989). In metazoan population studies, mitochondrial genes have played a prominent role due to the variability of the mitochondrial control region (Zischler et al., 1995). Their maternal inheritance and almost complete absence of recombination make the organelle genomes a unique tool for population biology. An interesting exception to the rule is the occasional recombination reported for fungal mitochondria (Saville et al., 1998). Several modes of inheritance have been reported from

different groups of fungi (Xu, 2005), but it is currently not known whether one of them occurs in the Glomeromycota.

In mitochondrial genomes of fungi including the one of *G. intraradices*, group I introns are widespread genetic elements (Lee & Young, 2009). They are capable of splicing by two sequential ester-transfer reactions (Cech, 1990) and many of them encode homing endonucleases (HEs). These enzymes are generally known for their role in proliferation of the introns they reside in. This "homing mechanism" involves cleavage of specific recognition sites (15-35 bp) in the exons with subsequent double-strand repair and insertion of the intron-containing allele at the respective site via recombination (reviewed by Chevalier & Stoddard, 2001). Several families of homing endonucleases are known based on conserved motifs, the most widespread being the LAGLIDADG family. The ORFs were thought to go through a "life cycle" consisting of fixation within a population, degeneration, intron loss and possible "re-colonization" (Goddard & Burt, 1999). Invasion of HEs in introns ensure their propagation and horizontal transfer of group I introns was reported to occur frequently between species of the same or of different kingdoms (Haugen et al., 2007). Alternatively, intron transfer can be promoted by reverse splicing. Contrary to homing, this process does not require a long recognition sequence but only few nucleotides (4–6 nt) that can pair bases with the internal guide sequence (IGS) of the intron (Cech, 1985; Woodson & Cech, 1989). Therefore, reverse splicing could give rise to transposition of introns into new genes.

Introns have been used as sensitive molecular markers in population studies (Neuveglise et al., 1997), because they tend to change very fast in evolution due to low selection pressure. On the other side, introns have been used to elucidate events of early evolution of land plants (Qiu et al., 1998; Cho, 1998). The view that some introns have been transferred from fungi to angiosperms very frequently (Vaughan et al., 1995; Cho et al., 1998; Sanchez-Puerta et al., 2008) was recently challenged and instead a history of ancient origin and frequent losses was suggested (Cusimano et al., 2008). Some introns are thought to be ancient and appear to have remained in the same position for millions of years (Qiu et al., 1998), some have apparently been transferred horizontally a relatively short time ago. In the light of these findings it would be interesting to better understand

the evolution of mtLSU introns in the Glomeromycota, which have turned out to be efficient molecular markers.

Previously available mtLSU sequences in the Glomeromycota were limited to the two relatively closely related morphospecies *G. intraradices* (*sensu lato*, see Material and Methods for details) and *G. proliferum*. The aim of the present study was to assess the evolutionary dynamics of the respective gene region from other glomeromycotan lineages in addition to their previously demonstrated discriminative power. In the focus was *Glomus* group A (Schwarzott et al., 2001), a monophyletic group which also contains *G. intraradices* and *G. proliferum*. Members of this clade represent a large part of glomeromycotan diversity and dominate almost all ecosystems studied so far using molecular identification methods. *Glomus mosseae*, *Glomus geosporum*, *Glomus caledonium*, *Glomus clarum* and *Glomus coronatum* were used as representatives of the second major clade within *Glomus* group A besides the *G. intraradices* clade. A species from *Glomus* group B, which constitutes the sister clade to *Glomus* group A, and two *Scutellospora* species from the more distantly related Gigasporaceae were used as outgroups. Using data from these species our objectives were to address the following questions:

- Does the lack of polymorphism in the mtLSU within fungal isolates also hold true for other species?
- Does the mtLSU represent a strain specific marker for other species?
- Do introns procure evidence about their evolutionary history and their potential spreading strategy possibly involving HEs?

III. Material and methods

1. Biological material

Spores of *G. coronatum*, *G. mosseae*, *G. caledonium*, *G. geosporum*, *G. clarum*, *Glomus* sp., *Scutellospora verrucosa* and *Scutellospora castanea* (Table 1) were harvested either from root organ cultures (ROCs) of *Daucus carota* (Bécard & Fortin, 1988) or from pot cultures (Table 1). Spores from ROCs were retrieved using 10 mM

sodium acetate-citrate buffer (pH 6.0) and washed in water (Doner & Bécard, 1991). Substrate from pot cultures (about 10 ml) was wet-sieved using a sieve cascade with openings of 1 mm combined with either 80 or 32µm. Organic matter from the sieves was suspended in 20 ml water, applied to a 70% (w:v) sucrose solution and centrifuged for 2 minutes at 820 *g* (Esch et al., 1994). The layer containing spores was rinsed in the bottom sieve transferred into Petri dishes before spores were placed in 1.5 ml tubes for DNA extraction.

Table 1: AMF used in this study.

Species	Isolate code	Host(s)	Location of origin
Glomus coronatum	ZTL	*Allium porrum*	Ghaziabad, India
Glomus mosseae	ISCB18	*Allium porrum*	Therwil, Switzerland
Glomus mosseae	ISCB13	*Allium porrum*	Biengen, Germany
Glomus mosseae	ISCB14	*Allium porrum*	Binningen, Switzerland
Glomus mosseae	BEG12	*Fragaria vesca*	Rothamsted, England
Glomus caledonium	BEG20	*Allium porrum, Hieracium pilosella, Plantago lanceolata*	Bedforshire, England
Glomus geosporum	BEG18	*Allium porrum, Hieracium pilosella, Plantago lanceolata*	Nenzlingen, Switzerland
Glomus clarum	BEG142	*Allium porrum, Hieracium pilosella, Plantago lanceolata*	Brazil
Glomus clarum	MUCL46238	*Daucus carota* (ROC)	Pinar del Rio, Cuba
Glomus sp.	ISCB34	*Hieracium pilosella, Plantago lanceolata*	Therwil, Switzerland
Scutellospora verrucosa	MN186	*Daucus carota* (ROC)	Central Kisa, Kenya
Scutellospora castanea	BEG01	*Allium porrum*	France

2. DNA extraction

DNA extracts were obtained by three different approaches: (i) DNA of *S. castanea* and all *Glomus* species was extracted from at least 20 spores using the DNeasy Plant Mini Kit (Qiagen, Hilden, Germany). (ii) Extracts of *Glomus* group A species were obtained from single spores as described by Redecker et al. (1997): each spore was crushed in 2 µl of NaOH (0.25 N). The extract was heated for 2 min at 95 °C before adding 1 µl of Tris-HCl (0.5 M, pH 8.0) and 2 µl of HCl (0.25 N) were added. DNA was incubated for 2 more min at 95°C and stored at -20°C. (iii) DNA of *S. verrucosa* and *S. castanea* was extracted from single spores and from hyphae using the approach of

(Redecker et al., 1997) and amplified with the GenomiPhi™ Amplification Kit (GE Healthcare).

3. Amplification of mtLSU

For the amplification of mtLSU DNA, primers were selected either from the literature or newly-designed using the software Primer Designer v.3.0 (Scientific & Educational Software, Cary, NC, USA). Based on sequences obtained from preliminary approaches (see Suppl. Doc. 1, Suppl. Fig. 1, Suppl. Table 1), a general improved nested PCR method was developed to amplify DNA from single or multiple spores of all species using the primers RNL-1/RNL-117 in the first step of the nested PCR and the primers RNL-29/RNL-118 in the second step of the nested PCR (Table 2, Fig 1). DNA from all species was successfully amplified. PCR reactions were carried out with a Phusion High-Fidelity DNA polymerase from Finnzymes (Bioconcept, Allschwil, Switzerland) in a total volume of 25 µl containing 1 µl of genomic DNA or 1 µl of water for the negative control, 1X Phusion HF buffer, 0.2 mM of dNTPs, 0.5 µM of each primer, 4X BSA, 3% DMSO, 0.02 U/µl of Phusion DNA polymerase. PCR products of the first step of the nested PCR were diluted 1:100 in water and used as template in the second step of the nested PCR. Cycling parameters were 30 seconds at 98°C, then 29 cycles of 10 seconds at 98°C, 20 seconds at 58.4°C, 2 minutes at 72°C and a final elongation of 10 minutes at 72°C. This approach was also tested on DNA extracted from roots colonized by *G. coronatum* using the DNeasy Plant Mini Kit (Suppl. Fig. 3).

Table 2: Primer sequences. The primers used in the general optimized nested PCR approach are shaded in grey. *, primers from Raab et al. (2005); **, primers from Börstler et al. (2008).

Primer name	Sequence (5'-3')	Primer name	Sequence (5'-3')
RNL- 1*	AGACCCGAARCCWRGTGATCT	RNL-103	CTGATTGCCCACCTTTGT
RNL-29**	TAATAAGACTGAACGGGTGT	RNL-104	CACACGGTACTACGTACCTA
RNL-11*	AAGGCAACACGCCAGCACTT	RNL-73	GCCTTATCACGAATCGTCTG
RNL-91	GCTCGTCTATACGTGCAACA	RNL-105	TGACGAATCGCCTTATCACG
RNL-92	AAGTGCTGGCGTGTTGCCTT	RNL-106	CAAGGCAGTCCTCTCCAGTG
RNL-93	TGCAGCTGGAAAAAGGAGTC	RNL-107	GCCACAGGTTGAATCCACTA
RNL-94	GCTGCTCCTTGAGGTCCTTA	RNL-108	ACGGAAGCTGAAGCTAGT
RNL-77	AGCCAACTCTATGGTTCAAT	RNL-109	CTCGGATCCTCAGAGACTAA
RNL-17*	CCATAGAGTTGGCTCTAACA	RNL-110	TGCTGAGTGACGTGCATTCT
RNL-95	ACGGACCTATTGCCCAATAC	RNL-111	AGCCGATCGAATGAGATACC
RNL-96	CGACTCATAACGGTGAATCC	RNL-112	CAGTCTAGGCACAAGGACTC
RNL-97	ATCGTAACACGTCGACACTC	RNL-113	GTGCCTAGACTGGAAGAGTA
RNL-98	GAGTGTCGACGTGTTACGAT	RNL-114	GCCTTGTKWTCACTCAGTAG
RNL-99	GGAGATTCTGCCTCTGTTTG	RNL-115	TTGGTTCCAGGCCTATCCTT
RNL-100	GCTAATCCTTCYGCTGTTAG	RNL-116	CCTGCTGATTGTCCAATCCT
RNL-101	AACTGTGGGGCAGTGTATGA	RNL-117	CTTCTGCTTYCGGCGAAGAG
RNL-102	GTCGACCTTTGCCCTCTTTG	RNL-118	TCARACCACTGAGCTGTTAC

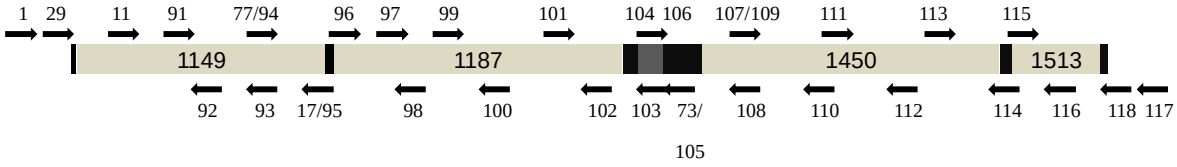


Fig. 1: Location and orientation of RNL primers. RNL-77, 17 and 73 were only used to sequence *G. clarum*. RNL-109-113 and RNL-107-108 were specific to *G. coronatum* and *Glomus* sp. respectively. Additional primers are shown in Suppl. Fig 1.

4. Amplification of nuclear rDNA

Nuclear rDNA amplification of each DNA extract was performed according to Redecker et al. (2000) using a *Taq* DNA Polymerase kit from GE Healthcare (Otelfingen, Switzerland), the universal eukaryote primers NS5/ITS4 for the first step of the nested PCRs and GLOM1310/ITS4i or GIGA1313/GIGA5.8R for the *Glomus* spp. and the *Scutellospora* spp. respectively for the second step of the nested PCRs.

5. Cloning of the PCR products and DNA sequencing

Blunt-ended PCR products obtained with Phusion polymerase were modified by an addition of an A overhang to the 3'-end by an incubation of 13 minutes at 72°C with *Taq* polymerase, 2mM $MgCl_2$ and 0.125 mM dATP. All PCR products were purified using the High Pure Kit (Hoffmann LaRoche, Basel, Switzerland). Cloning was performed according to the manufacturer's instructions of the pGEM® Vector System I kit (Promega/Catalys, Wallisellen, Switzerland). Cloned PCR products were re-amplified with the vector primers M13f (GTAAAACGACGGCCAGTG) and M13r (GGAAACAGCTATGACCATG). After purification of the PCR products, DNA was resuspended in 30 µl of elution buffer and stored at -20°C until use.

The ABI PRISM Big Dye Terminator v.3.1 Cycle Sequencing Ready Reaction kit® (Applied Biosystems, Foster City, CA) was used to perform the sequencing PCR in both directions. Samples were run on an ABI prism 310 capillary sequencer (Applied Biosystems) or on an ABI3130xl capillary sequencer.

6. Sequence and phylogenetic analyses

The sequences obtained were edited and compiled using Sequence Navigator software (version 1.0.1, Applied Biosystems). In case of sequences obtained from the same DNA extract but with different primer sets, a concatenated sequence was created based on evidence of mtLSU homogeneity and only when a significant overlap (≥ 190 bp) allowed excluding ambiguity. Sequences were manually aligned to previously published sequences (Raab et al., 2005; Börstler et al., 2008) with Bioedit (Hall, 1999) or PAUP*4.0b10 (Swofford, 2001). Sequences were deposited in the EMBL database under the accession numbers FN377857 to FN377867 and FN377601 for mitochondrial sequences and FN423686 to FN423707 and FN423502-FN423502 for nuclear sequences. Phylogenetic trees were generated by distance, parsimony and maximum likelihood criteria using PAUP*4b10 (Swofford, 2001). Modeltest 3.5 (Posada, 2004) was used to estimate maximum likelihood models and parameters and MrBayes 3.1.1 (Ronquist & Huelsenbeck, 2003) to perform Bayesian analyses. For parsimony analyses each gap of three bases or more was coded by inserting a binary character (1) in introns.

Unless indicated otherwise, Bayesian analyses were run for 1 000 000 generations with a "burnin" of 10%. The stationarity of the Markov chains was verified using the command "sump". For protein phylogenies, a mixed model was used that integrates different protein sequence evolution models.

Introns were identified by the software RNAweasel (Lang et al., 2007). Search of ORFs was performed by means of the programme NEB Cutter Version 2.0 (Vincze et al., 2003) with the "Mold, Protozoan, and Coelenterate Mitochondrial Code and the Mycoplasma/Spiroplasma Code" (transl_table=4). Detected ORFs were translated, submitted to the BLASTp network server (NCBI; www.ncbi.nlm.nih.gov) and conserved domains were detected using the Conserved Domain Database (CDD; Marchler-Bauer et al., 2009). Statistical modeling of LAGLIDADG allowed the observation of single versus double-motif HEs as well as intact versus disrupted conserved domains (Belfort & Roberts, 1997; Dalgaard et al., 1997; Heath et al., 1997). Additionally, intronic sequences were translated in six reading frames (http://www.expasy.ch/tools/dna.html) to observe any putative LAGLIDADG domains that were not encoded by ORFs.The rates of synonymous (dS) and non-synonymous (dN) substitutions were computed using the Synonymous Non-synonymous Analysis Program (Korber, 2000).

For simplicity, we use the species name *G. intraradices* in the broad sense throughout this study for isolates morphologically identified as this species. It has been shown that within this definition several distinct divergent lineages can be distinguished by nuclear-encoded rDNA sequences (e.g. Hijri et al., 2006) and mtLSU (Börstler et al., 2008). Recently, it was shown that the lineage comprising most isolates, environmental sequences, and in particular the strain currently used for sequencing the genome of *G. intraradices*, is distinct from the type strain of the species (FL 208; Stockinger et al., 2009). Consequently, the majority of isolates and environmental sequences will have to be renamed in the future. Until a generally accepted epithet for this enormously widespread fungal clade is at hand, we will refer to this clade as *G. intraradices* GLOM-A1 in concordance with the phylotype nomenclature of our previous field studies (e.g. Hijri et al., 2006).

IV. Results

1. Primer design and PCR

So far, primers of Raab et al. (2005) and Börstler et al. (2008) were designed to amplify *G. proliferum* and *G. intraradices* but no set of primers was available allowing the detection of other taxa. Primers reported previously and newly designed primers (Suppl. Table 1) were used to amplify parts of the mtLSU from five species of *Glomus* group A (*G. clarum*, *G. geosporum*, *G. mosseae*, *G. caledonium*, *G. coronatum*), one species from *Glomus* group B and two species of the more distantly related Gigasporaceae. From the obtained sequences, a new set of primers (RNL-1/RNL-117 and RNL-29/RNL-118) was designed allowing detection of all samples. The primer pairs RNL-1/RNL-117 and RNL-29/RNL-118 were successfully used in a nested PCR approach, yielding products of sizes ranging from 538 to 4003 bp (Suppl. Fig. 2, Table 3).

All PCR products could in principle be directly sequenced after PCR, but most amplicons were sequenced after cloning because this proved to be more efficient as numerous sequencing reactions were required to cover the long DNA fragments. Nuclear-encoded rRNA was amplified from the same biological samples and analyzed phylogenetically as a control for species identity (Suppl. Fig. 4, Suppl. Fig. 5).

The general mtLSU primer set was specific enough to be used for amplification of fungal DNA from colonized roots, as shown by successful amplification of the expected fragment from *Allium porrum* colonized roots (Suppl. Fig. 3).

Table 3: Exon/intron structure of the mtLSU sequences between the priming sites RNL-29/RNL-118. *G. intraradices* JJ291and *G. proliferum* DAOM226389 were included as references. *Numbers denote positions in the *G. intraradices* exon FACE494 (Lee & Young, 2009). Each table cell lists intron length and intron type.

	Introns				Exons	
Species – Accession numbers – (Isolates)	Position 1149* (Length, bp) – intron group	Position 1187* (Length, bp) – intron group	Position 1450* (Length, bp) – intron group	Position 1513* (Length, bp) – intron group	Complete length/parts separated by introns (bp)	Total length (bp)
G. proliferum (DAOM226389)	No	No	No	No	386/39, 261, 62, 24	386
G. intraradices (JJ291)	1056 – IA3	No	No	401 – IB	386/39, 261, 62, 24	1843
G. clarum FN377859 (BEG142) FN377601(MUCL46238)	1057 – IA3	No	No	No	384/39, 259, 62, 24	1441
G. geosporum FN377861 (BEG18)	1133 – IA3	No	No	272 – IB	497/39, 373, 61, 24	1902
G. mosseae FN377862 (ISCB18) FN377863 (BEG12) FN377864 (ISCB13) FN377865 (ISCB14)	1136 – IA3	1378 (1379 for ISCB 18) – IA3	No	272 – IB	496/39, 372, 61, 24	3282
G. caledonium FN377857 (BEG20)	1101 – IA3	1374 – IA3	No	272 – IB	496/ 39, 372, 61, 24	3243
G. coronatum FN377860 (ZTL)	1181 – IA3	606 – IA3	1448 – I derived	272 – IB	496/39, 372, 61, 24	4003
Glomus sp. FN377858 (ISCB34)	No	No	1045 – IB	215 – IB	417/39, 293, 61, 24	1677
S. verrucosa FN377866 (MN186)	No	No	No	No	538/39, 414, 61, 24	538
S. castanea FN377867 (BEG01)	No	No	No	No	545/39, 421, 61, 24	545

2. Exon/intron structure

The broad size range of the amplicons was mainly caused by introns inserted in four locations (Fig. 2, Table 3). The two *Scutellospora* spp. did not contain any intron, whereas *G. coronatum* on the other extreme contained all of them. All introns were identified to be group 1 introns except the *G. coronatum* intron position 1450 which was characterized as "group I derived". Several species contain short insertions (44 to 165 bp) absent in others, which in the absence of evidence defining them as introns are however regarded as "expansion segments". The occurrence of such expansion segments or

nucleotide extensions in the mtLSU was firstly reported in Basidiomycota (*Agrocybe aegerita*), Ascomycota (*S. cerevisiae*) and also in algae (*Prototheca whickerhamii*), (Sor & Fukuhara, 1983; Wolff et al., 1993; Gonzalez et al., 1999) possibly constituting parts of the exon highly variable in length. The evolution of these fragments is discussed by Hancock et al. (1988) and Yokoyama et al. (2008). They were excluded from phylogenetic analyses of the exon because they were difficult to align and decreased the support of the phylogenetic trees.

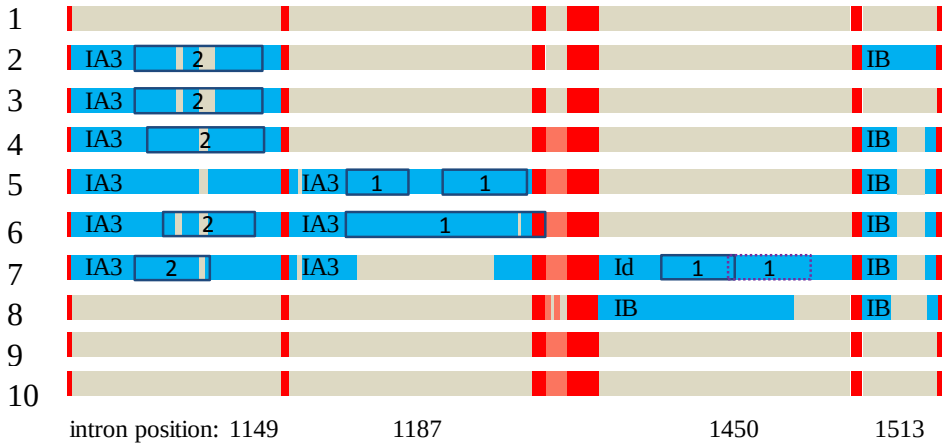


Fig. 2: Schematic sequence alignment of the mitochondrial rRNA large subunit (mtLSU) gene region for the following isolates of different species: 1. *G. proliferum* DAOM226389 – 2. *G. intraradices* JJ291 – 3. *G. clarum* BEG142 / MUCL46238 – 4. *G. geosporum* BEG18 – 5. *G. mosseae* ISCB18 – 6. *G. caledonium* BEG20 – 7. *G. coronatum* ZTL – 8. *Glomus* sp. ISCB34 – 9. *S. verrucosa* MN186 – 10. *S. castanea* BEG01. Exons are represented in red (), introns in blue (), gaps in grey () and insertions in pink (). The intron type is indicated by IA3, Id for group I derived and IB. Open reading frames appear in rectangular boxes with the number corresponding of the putative LAGLIDADG type. Diagram approximately to scale.

3. Intraspecies sequence homogeneity

Four isolates of *G. mosseae* originating from Switzerland (two different sites), Germany and United Kingdom were analysed including the strain BEG12 widely used in previous studies (e.g. Gamalero et al., 2004; Repetto et al., 2007; Pivato et al., 2009). To further verify sequence polymorphism of the mtLSU within and between isolates of *G. mosseae*, sequences were amplified and sequenced from 3 up to 7 spores for each isolate. No sequence variation was detected between the clones due to the fidelity of the proof-

reading DNA polymerase used: the error rate of Phusion is determined to be 4.4×10^{-7} according to the manufacturer. Sequences of *G. mosseae* ISCB13 and BEG12 originating from Germany and United Kingdom respectively were identical. Sequences of *G. mosseae* ISCB18 and ISCB14 both originating from Switzerland but at different sites were 99.8% similar. Sequences of *G. mosseae* ISCB18 and *G. mosseae* ISCB14 were 99.9% and 99.9% respectively similar to *G. mosseae* ISCB13 and BEG12. On average, the percentage of similarity between these four isolates was 99.9%. Two isolates of *G. clarum* BEG142 and MUCL46238 originating from Brazil and Cuba, respectively, were also analyzed. They also showed a remarkably high similarity of 99.9%. In consequence, mtLSU sequences were considered as being identical between different isolates of *G. mosseae* and *G. clarum*.

4. Exon phylogeny

The relatively short exon regions yielded a phylogeny largely consistent with the one obtained using nrDNA (Fig. 3). Gigasporaceae, *Glomus* group A and *Glomus* group B grouped as expected and while they could be clearly assigned to the Glomeromycota, they showed considerable genetic distance from each other. The well-established sister group relationship between the *G. mosseae* and the *G. intraradices* clades was recovered to different extents by different phylogenetic analyses. In parsimony- and maximum likelihood-based and Bayesian analyses the *G. mosseae* clade was not resolved as monophyletic group, which is most likely due to the low phylogenetic signal (in particular parsimony-informative characters) differentiating this clade. The isolate CA502 showed a closer affinity to *G. clarum* than in ITS-based phylogenies, where it groups within the GLOM-A1 clade.

Overall, the polymorphism in the exon region studied was relatively low between closely related species, with no differences at all between *G. mosseae/coronatum* and 99.2% similarity between *S. verrucosa/S. castanea*. In the 18S n-rDNA based phylogeny these two pairs of species are very closely related sister taxa (Suppl. Fig. 4).

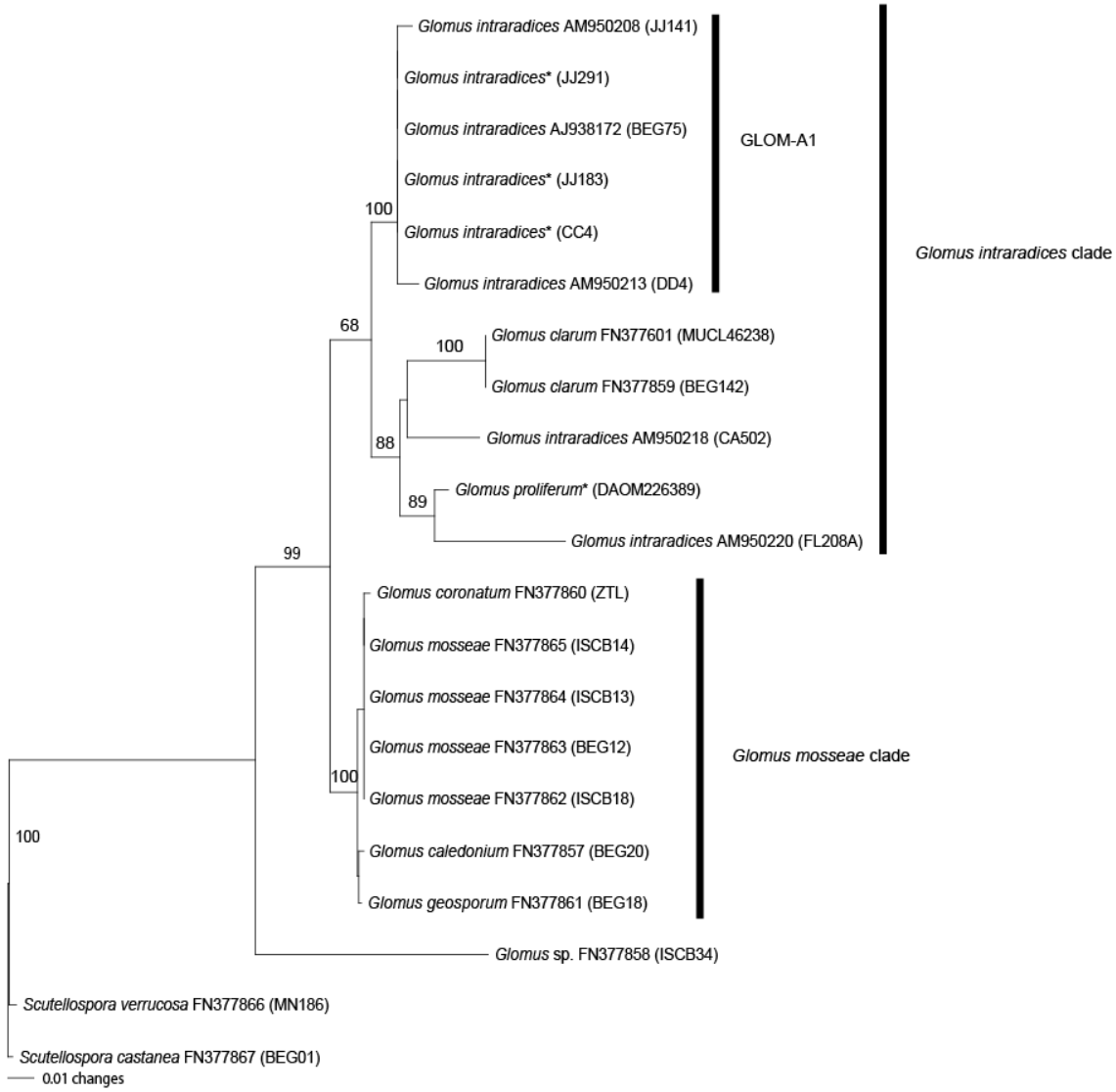


Fig. 3: Neighbor-joining tree of Glomeromycota based on 361 bp of exon mtLSU, with *Scutellospora* spp. used as outgroup. Values on the nodes indicate neighbor-joining bootstrap values from 1000 replicates. After species name are accession numbers and isolates (in brackets). *, consensus sequences obtained from AM950210, AM950211 and AM950212 for *G. intraradices* JJ183; AJ973189, AJ973190 and AJ973191 for *G. intraradices* JJ291; AM950204, AM950205 and AM950206 for *G. intraradices* CC4; AM040980, AM040981 and AM040982 for *G. proliferum*.

5. Evolution of ORFs and conserved domains of the LAGLIDADG in introns

Intron 1149: evidence for intraphylum horizontal transfer

Within intron 1149, *G. geosporum* and *G. clarum* displayed ORFs for LAGLIDADG HEs with one conserved domain, as for *G. intraradices* isolate JJ291 even though in the latter, the conserved domain was disrupted. In *G. intraradices* isolates JJ141 and JJ183, the entire conserved domain was encoded in two separate putative ORFs (Börstler et al., 2008). In *G. coronatum* and *G. caledonium* the ORFs were shorter compared to *G. geosporum* and *G. clarum*, because some parts of the domain were shifted in frame, but the overall intron length was similar (Fig. 2). In *G. mosseae*, no ORF was detected, but interestingly some parts of the conserved motif were still present in the respective region. In three isolates of *G. intraradices* (DD-4, CA502, FL208A), neither ORFs nor conserved domains were present and the size of intron 1149 in these isolates (425 to 662 bp) was substantially shorter, because the sequences coding for the ORFs in the other isolates were lacking.

The ratio of synonymous (dS) to non-synonymous (dN) changes (Suppl. Table 3) showed an overall dS/dN value of 1.6217, indicating a negative selection on the HE ORFs to maintain the protein activity. The dS/dN ratio was higher in functional (2.0119) compared to disrupted (1.3913) endonuclease domains. This can be interpreted as a faster degeneration of HE ORFs non-encoding putative active LAGLIDADG. The shortened *G. coronatum* HE ORF even showed stronger signs of degeneration (dS/dN =0.8287).

Therefore, within the *G. mosseae* clade, there appears to be a progression from a large ORF containing a LAGLIDADG domain towards an apparent partial or entire loss of the ORF by losing start codons and/or introducing stop codons. Interestingly, even in *G. mosseae*, which does not possess an ORF, the degradation of the sequence is limited.

The phylogeny of the LAGLIDADG ORFs and the remaining parts of the intron was in congruence which favours the hypothesis of transfer of both endonuclease and intron 1149 as a unit (Fig. 4, Suppl. Fig. 6). Some isolates of *G. intraradices* grouped differently from the exon phylogeny. In particular, the intron of JJ183 showed a striking similarity to

the *G. clarum* isolates, sharing several characteristic motifs. This is evidence for horizontal transfer of the intron, either from *G. intraradices* JJ183 to *G. clarum* or vice versa. Similarly, *G. intraradices* JJ141 groups with the *G. mosseae* clade.

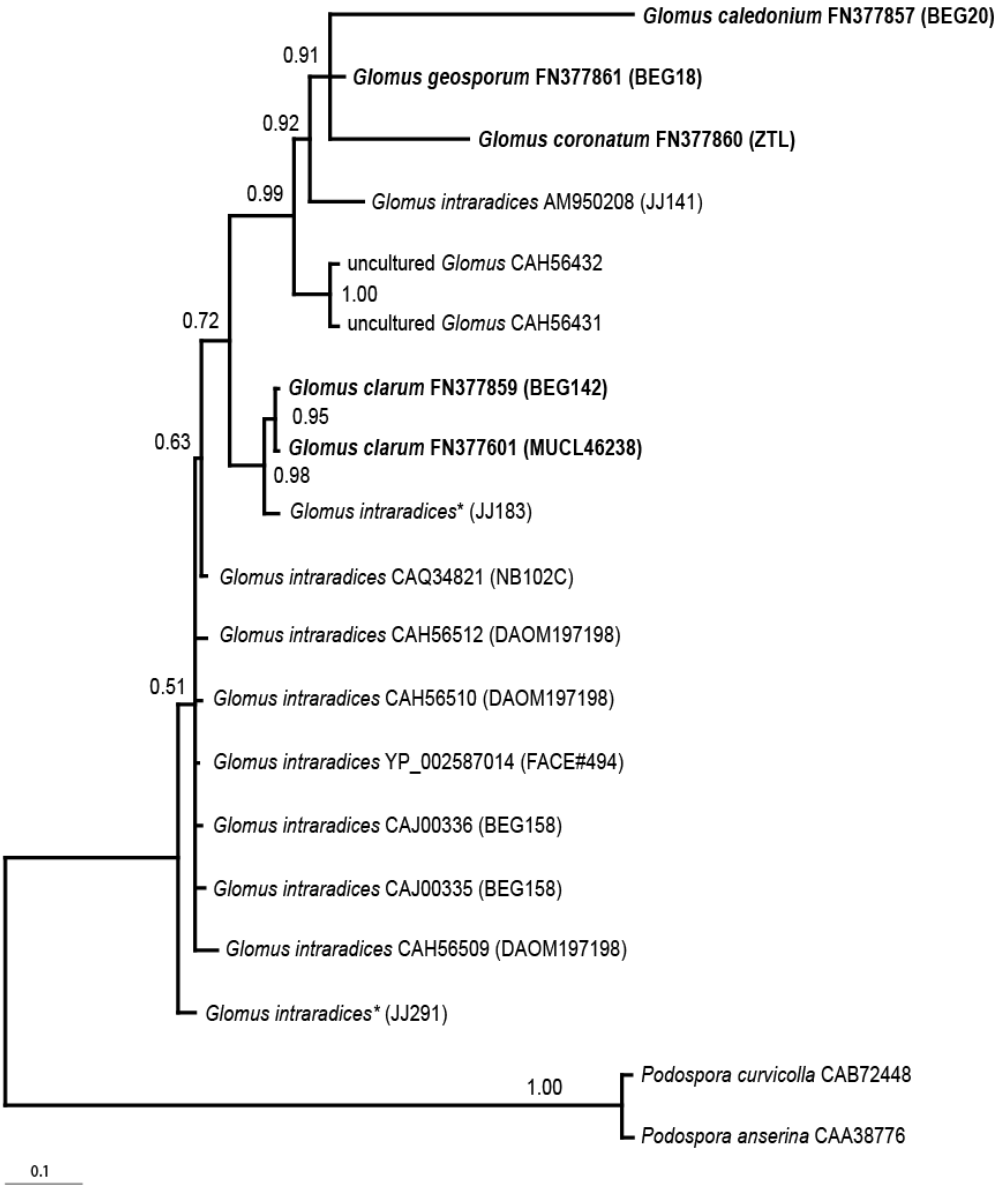


Fig. 4: Phylogeny of the putative LAGLIDADG type 2 proteins encoded by intron 1149 ORFs in the Glomeromycota, with Ascomycota as outgroup. The phylogenetic tree was generated based on an alignment of 267 amino acids. Values on the nodes indicate Bayesian posterior probabilities. Labels include species names, accession numbers and isolate codes (in brackets). *, consensus sequences obtained from AM950210, AM950211 and AM950212 for *G. intraradices* JJ183; AJ973189, AJ973190 and AJ973191 for *G. intraradices* JJ291.

Intron 1187: secondary split of double-motif LAGLIDAG ORFs

Intron 1187 was only found in three species, therefore no phylogenetic analysis was conducted. Two species contained putative LAGLIDADG type 1 ORFs in this intron: *G. caledonium* possessed two copies of the conserved amino acid motif LAGLIDADG whereas *G. mosseae* had two distinct ORFs for each conserved LAGLIDADG motif. This split was due to a stop codon arising from a deletion of 5 bp. In *G. coronatum* this intron was present but its size was reduced from 1378 bp to 606 bp. Here, the start codon and some other remaining sequences from the LAGLIDADG ORF were still present, indicating a large part of the ORF was lost, and providing conclusive evidence for the direction of this evolutionary process. In *G. geosporum*, which is closely related to *G. caledonium*, the intron was completely absent.

In order to elucidate the phylogenetic relationships of HEGs (HE genes) of Glomeromycota and other taxa with single or double conserved LAGLIDADG domains, the phylogenetic analysis of the HEGs (Fig. 5) was conducted on each conserved domain: for that purpose the N-terminal and C-terminal domains of the two conserved domains of double-motif LAGLIDADG were split, aligned to each other and to the two ORFs from *G. mosseae*, which clearly represent homologs of the two domains in *G. caledonium*. Overlapping ORFs of HEs found in *G. coronatum* in intron 1450 were also included in this analysis.

Single-domain LAGLIDADG from organelles of algae and other organisms were used as outgroup since it has been shown that double-motif LAGLIDADGs have evolved by duplication of single-motif proteins (Belfort & Roberts, 1997). A monophyletic group almost exclusively containing fungal two-domain LAGLIDADG was recovered, which was well-supported in Bayesian analysis (posterior probability=1.00). Single-domain LAGLIDADG ORFs from *G. mosseae* and *G. coronatum* as well as from the green alga *Chlorella* in this clade were the only exception we detected. The sequence from the alga grouped rather basally. Interestingly, two overlapping single-domain ORFs from intron 1450 of *G. coronatum* also grouped basally in this clade. Their origin from single or double-domain LAGLIDADG remains unclear.

Besides two clades formed by mtLSU-hosted N-terminal and C-terminal domains of fungal HEGs, respectively, two more monophyletic groups were found, representing the domains of LAGLIDADG inserted in other fungal mitochondrial genes, mainly NADH. The N-terminal domains from mtLSU and NADH formed a clade highly supported by posterior probabilities, indicating that they share a common ancestor, whereas a clade of the C-terminal domains was not supported, but neither could be rejected. The tree indicates that the two single motifs of *G. mosseae* are clearly derived from double-motif ancestors.

This phylogenetic analysis indicates that mtLSU glomeromycotan HEGs spread to other mitochondrial target genes like NADH of other fungi, mainly Ascomycetes. Interestingly, the clade appears to be restricted to fungi, suggesting that the HEGs spread within the true fungi or even coevolved with them.

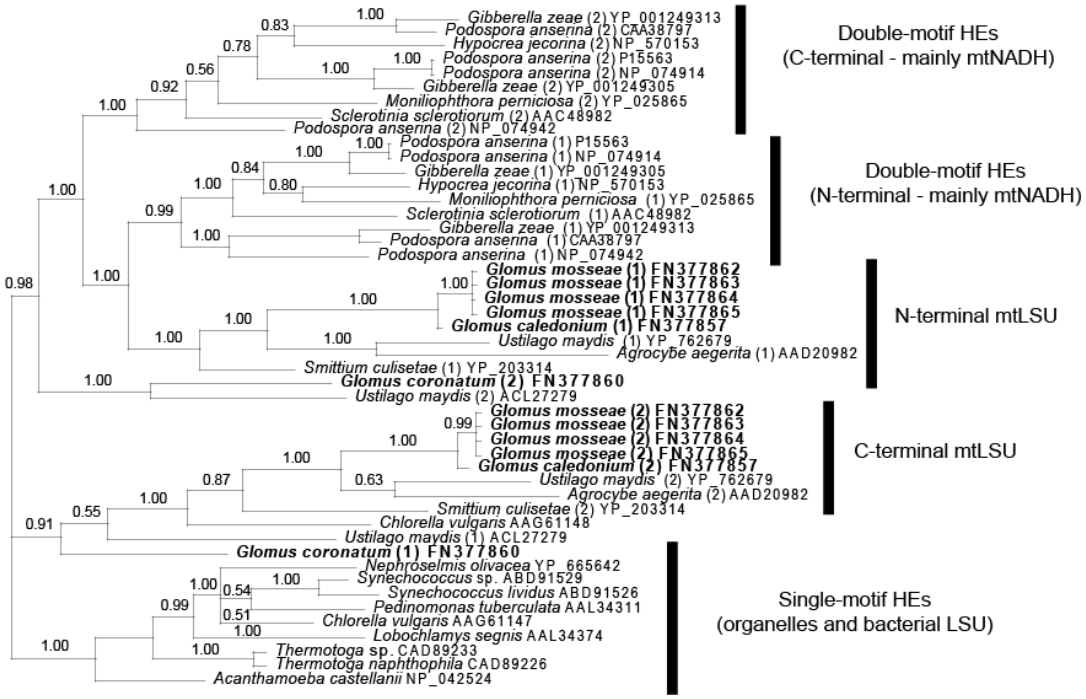


Fig. 5: Phylogenetic tree of putative LAGLIDADG type 1 proteins, based on an alignment of 143 amino acids. Values on the nodes indicate Bayesian posterior probabilities. LAGLIDADG conserved domains were treated separately in the alignment. Numbers in brackets designate the N-terminal domain (1) or the C-terminal domain (2) for double-motif HEs, or in the case of *G. mosseae* and *G. coronatum* for distinct ORFs in the same intron.

Intron 1450: two different intron types in the same insertion site

As mentioned above, *G. coronatum* displayed overlapping genes of LAGLIDADG type 1 HEs, orientated in the same direction, which could be attributed to a duplication event. This process may have been the same as the one originally leading to the domain duplication in LAGLIDADG type 1 gene of *G. caledonium* by duplication of overlapping sequence regions and their elimination (Keese & Gibbs, 1992).

Interestingly, the *Glomus* sp. ISCB34 intron 1450 was not related to the *G. coronatum* intron in the same position, they did not even belong to the same subtypes (Table 3). No conserved domains or ORFs were recognizable in any of the six frames translated from the intronic sequence.

Intron 1513: vertically inherited from the ancestor of Glomus groups A and B

Translation of the intronic sequence in six reading frames did not show any conserved domains of endonucleases similarly to the *Glomus* sp. ISCB34 intron 1450, which belongs to the same subtype (IB). The phylogeny of this intron (Fig. 6) showed a close similarity to the exon phylogeny (Fig. 3). Together with the fact that it was also present in a member of *Glomus* group B, this suggests that it was present in the common ancestor of *Glomus* groups A and B and was inherited in a rather conservative manner throughout the evolutionary history of these groups with occasional losses.

There is some evidence that reverse splicing (Bhattacharya et al., 2005) could play a role in the spread of introns of this group, which did not show hints for endonuclease ORFs: a conserved motif (5'-TCAAAT-3') was found at the insertion site of the intron 1450 and was also present just downstream of the 5' splice site of the intron 1513.

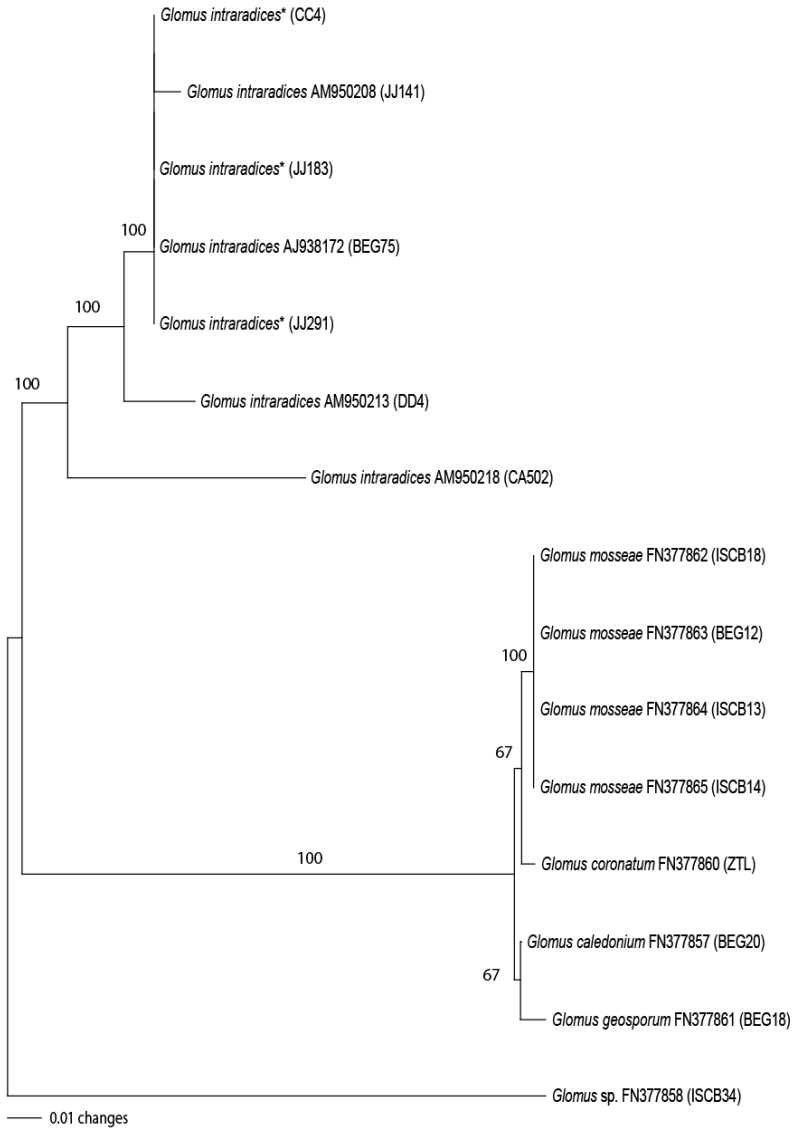


Fig. 6: Phylogeny of Glomeromycota based on 261 bp of intron 1513 sequences. The phylogenetic tree was generated by a heuristic search under the parsimony criterion. Numbers at the node indicate bootstrap values for 1000 replicates. Labels include species name, accession numbers and isolate codes (in brackets). *, consensus sequences made from AM950210, AM950211 and AM950212 for *G. intraradices* JJ183; AJ973189, AJ973190 and AJ973191 for *G. intraradices* JJ291; AM950204, AM950205 and AM950206 for *G. intraradices* CC4.

V. Discussion

(Raab et al., 2005) showed that in strong contrast to many nuclear-encoded genes, in particular rDNA, mtLSU in the two glomeromycotan species *G. intraradices* and *G. proliferum* was homogeneous within the same strain. Lee and Young (2009) confirmed this finding for the complete mitochondrial genome of *G. intraradices*. Here we show that this genetic homogeneity of mtLSU is also a general feature of other taxa of AMF.

While not showing within-strain mtLSU polymorphisms, many isolates of *G. intraradices* differed strongly with regard to their mtLSU introns (Börstler et al., 2008), allowing insights into the population structure of this ubiquitous species, which has also gained importance as model organism because its genome is being sequenced (Martin et al., 2008). Species definition of *G. intraradices* is problematic and within the current morphology-based circumscription, several genetic lineages were identified that may eventually correspond to species. In particular, it was demonstrated that the ubiquitous lineage containing the strain used for genome sequencing (GLOM A-1) is distinct from the type strain (Stockinger et al., 2009). However, as some of these isolates of different mtLSU haplotypes were shown to fuse hyphae and exchange genetic markers, these haplotypes do not necessarily correspond to cryptic species (Croll et al., 2009).

In contrast to the high level of polymorphism among *G. intraradices* isolates, we show that two *G. clarum* isolates and four *G. mosseae* isolates from different locations did not differ at all in the respective sequence region. We cannot exclude that by analyzing a larger number of isolates some polymorphisms would be eventually obtained, but the low level of divergence even of the related morphospecies *G. coronatum*, *G. geosporum* and *G. caledonium* does not make this kind of search promising. The low level of genetic polymorphism in *G. mosseae* is in agreement with other authors' findings (Rosendahl & Matzen, 2008), that this species does not show geographical structure across continents (Giovannetti et al., 2003). The biological reason for this apparent discrepancy between the *G. intraradices* and the *G. mosseae* groups currently remains unclear but it will be interesting to address this on a broader scale by analyzing additional families of the Glomeromycota.

As the species in the *G. mosseae* clade, the two *Scutellospora* species are relatively closely related, as judged by their partial 18S n-rDNA sequences. However, in their mtLSU exon short but distinct sequence signatures were present distinguishing them.

The introns found in the different *Glomus* spp. provide evidence for a variety of processes of intron and HEG evolution. This is the first time that these processes are addressed in the phylum Glomeromycota.

Among the introns, there is a wide gamut of the different degrees of conservation. Intron 1513 has apparently been inherited strictly vertically from the ancestors of *Glomus* groups A and B, and its phylogeny correlates well to the exon phylogeny. This group IB intron was already shown to be relatively stable in *G. intraradices* (Börstler et al., 2008).

The phylogeny of intron 1149 was not completely concordant with what would be expected from the exon and n-rDNA. In particular the striking similarity of the corresponding sequences from *G. intraradices* JJ183 and *G. clarum* indicates occasional horizontal transfer. However, all evidence was in favor of a transfer of the intron together with its HEG, which is in agreement with some findings of Haugen and Bhattacharya (2004) even though these authors also reported intron-independent mobility of HEGs (Haugen et al., 2004). In this intron, we can see a gradient of purifying selection between putatively active endonucleases and putatively non-functional sequences. The overall negative selection in intron 1149 may have played a role in maintaining LAGLIDADG genes to ensure their spreading to intron-less individuals. Evolutionary pressure still exists in disrupted endonuclease domains of *G. intraradices* isolates JJ141 and JJ183 as well as in *G. caledonium* but is relatively low in the putative non-functional ORF of *G. coronatum*, suggesting that when homing activity ceases, the degeneration of HE accelerates until its total suppression. The presence, maintenance and degeneration of LAGLIDADG ORFs observed in this study are highly compatible to the stages in theory of "life cycle" described by Goddard & Burt (1999). Interestingly, even in *G. mosseae*, which does not seem to possess an intact ORF, the degradation of the sequence is limited. We can speculate that purifying selection might occur on parts of introns required for self-splicing.

In the *G. mosseae* clade, we found different stages of the degradation of the HEG-containing intron 1187: intact ORF in *G. caledonium*, the ORF split into two parts in *G. mosseae* and the ORF lost from the intron in *G. coronatum*. *G. geosporum* which lacks this intron may represent an ancestral state before the intron was inserted. However, considering the established n-rDNA phylogeny of these species which places *G. geosporum/G. caledonium* as sister group to *G. mosseae* and *G. coronatum* (Schwarzott et al., 2001), it appears more likely that the common ancestor of the clade possessed the intron and *G. geosporum* lost it completely. It will be intriguing to address in future studies whether different stages of the "homing cycle" can be detected within populations of each of these species as demonstrated for other fungi by Reeb et al. (2007). In *G. intraradices* isolates, different stages of degradation of HEG ORFs are known, in *G. mosseae* on the other hand we could not find any polymorphisms so far in the introns.

Haugen and Bhattacharya (2004) analyzed the evolutionary relationships of HEGs and postulated gene duplication through which double-motif HEGs arose from single-motif HEGs. Interestingly, these authors found exclusively fungal sequences in the monophyletic clade of double-motif HEs, a fact that they did not discuss further. This "clade 1" corresponds to the clade of double-motif HEs in Fig. 5. The strict predominance of fungal sequences in this clade indicates that these HEs have evolved and may even have originated within the true fungi.

In contrast to Haugen and Bhattacharya (2004), we also included HEGs in the dataset which are hosted in other mitochondrial genes than LSU, for instance NADH3. The phylogenetic analyses (Fig. 5) suggest that the spread towards other genes occurred after the domain duplication. The fact that only the N-terminal domains in mtLSU and other genes are supported as a monophyletic group may indicate that this target change occurred relatively quickly after the duplication. This interpretation would be compatible with the hypothesis that duplication facilitates the spread of the HE to a broader range of target sequences, because single domain HEs assembling into homodimers are mainly limited to palindromic target sites (reviewed by Gimble, 2000).

Interestingly, the two overlapping single-motif ORFs in intron 1450 of *G. coronatum* also belong to this major clade, but evolved separately. It is intriguing to speculate that

these ORFs are the remnants of another lineage of double-motif HEGs that by a frameshift fragmented in a way reminiscent of the HEG in intron 1187 in *G. mosseae*, but with non-coding sequences between the ORFs. Alternatively, the observed overlapping could represent an early stage of the fusion of two HEGs into a two-domain ORF.

In contrast to other studies claiming rampant interkingdom intron transfer (Cho & Palmer, 1999), which however were challenged by other authors (Cusimano et al., 2008), most introns in the DNA region we studied appear to be predominantly spreading among the true fungi. This is in agreement with the results of Haugen and Bhattacharya (2004), who demonstrated extensive change of target sites within the same target gene (LSU), but exclusively within true fungi. However, intron transposition between mycorrhizal fungi and plant hosts seems possible as suggested by Lang & Hijri (2009).

HEs have recently received considerable attention in biotechnology as model systems for the engineering of enzyme activity (Silva et al., 2006). This is possible because active domains and DNA target sites are well-characterized and can be manipulated. Our data provide insights into the in situ dynamics of HEGs in a natural system, demonstrating major stages of the "homing cycle" (Chevalier & Stoddard, 2001). By providing an evolutionary snapshot of LAGLIDADG HEGs in a group of "lower fungi", which have been strongly underrepresented in the databases, this study will contribute to a better understanding of the evolution of these ubiquitous genes.

VI. Acknowledgments

We would like to acknowledge the Swiss National Science Foundation for funding this research (grants 3100A109466 and 3100A129466 to D. R.), Thomas Boller and Andres Wiemken at the Botanical Institute of the University of Basel for continuing support, Mathimaran Natarajan for providing DNA of *S. verrucosa*, and all the members of the “mycorrhiza” group at the Botanical Institute for helpful discussions and the technical staff for their assistance. We also would like to thank Hannes Gamper and Thomas Boller for helpful comments on our manuscript.

VII. Supplementary files

Suppl. Doc. 1: Nested PCR approaches.

The mt-LSU rRNA gene was first amplified through nested PCR with the *Taq* DNA Polymerase Kit (GE Healthcare, Otelfingen, Switzerland) using previously published or newly designed primers (Table 1, Table 2). PCR reactions were performed in a total volume of 25 µl containing 1 µl of genomic DNA or 1 µl of water for the negative control, 1x PCR buffer, 2 mM of $MgCl_2$, 0.25 mM of dNTPs, 0.5 µM of each primer, 0.05 U/µl of *Taq* polymerase. The PCR program consisted of a denaturation step of 3 min at 95°C, followed by 34 cycles of 1 min at 95°C, 1 min at 51 or 56°C for the first and the second step of the nested PCR respectively, and 4 min at 72°C. The PCR ended with a final elongation of 10 or 5 min, for the first and the second step of the nested PCR respectively, at 72°C. The polymerase used and the PCR conditions are referred to approach 1 (Table 4).

In order to increase PCR yield and fidelity, PCR reactions were carried out with a Phusion High-Fidelity DNA polymerase from Finnzymes (Bioconcept, Allschwil, Switzerland) in a total volume of 25 µl containing 1 µl of genomic DNA or 1 µl of water for the negative control, 1x Phusion HF buffer, 0.2 mM of dNTPs, 0.5 µM of each primer, 4x BSA, 3% DMSO, 0.02 U/µl of Phusion DNA polymerase. New primers were designed (Table 2). Cycling conditions were changed as following: denaturation of 30 s at 98°C, then 34 cycles of 10 s at 98°C, 30 s at 51°C, 2 min at 72°C. The PCR ended with a final elongation at 72°C of 10 min. In the case of *G. clarum*, the PCR conditions were 30 s at 98°C, then 37 cycles of 10 s at 98°C, 20 s at 58°C, 2 min at 72°C. The PCR ended with a final elongation at 72°C of 10 min. This is referred to approach 2 (Table 4).

At this stage, DNA was amplified with the primers (i) RNL3/RNL15 for Gigasporaceae (ii) RNL29/RNL15 for *G. geosporum* (iii) RNL1/RNL5 for *G. clarum* (see also Börstler *et al.* 2010) (iv) RNL1/RNL98 – RNL96/RNL116 – RNL129/ RNL139 or RNL128/RNL141* for *G. caledonium*, *G. mosseae* and *G. coronatum**.

The final approach described in Material and Methods section is referred to approach 3.

Suppl. Table 1: Additional primers used in the two first approches, which are not required in approach 3. *, primers from (Raab et al., 2005); **, primers from Börstler et al. (2008).

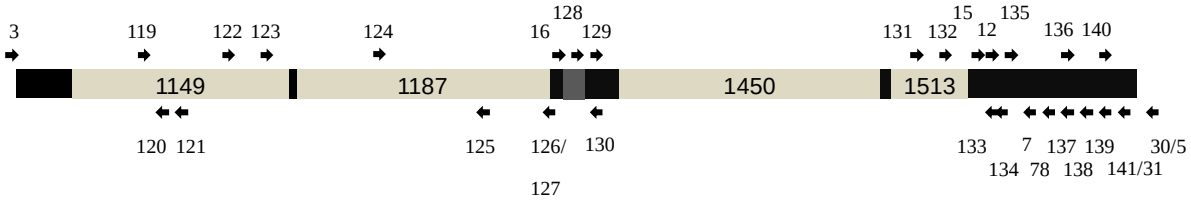
Primer name	Sequence (5'-3')	Primer name	Sequence (5'-3')
RNL- 3*	TGCATMATGGGTCAGCGAGT	RNL-15*	CTGAGCTGTTACGCTATC
RNL-119	TCCTCGTAGGAGGAATGT	RNL-12*	GATAGCGTAACAGCTCAGTG
RNL-120	TGTTGCACGTATAGACGAGC	RNL-133	CAGACCGCTGAGCTGTTACG
RNL-121	CAGACGACACGTTGCAGCTG	RNL-134	GCGGAAGACCAGCTAATC
RNL-122	TTGGGCCGAGGGGGGTATAA	RNL-135	AGGCGAACCTTCGCTTGTAA
RNL-123	GTAAGGTCTATCGGGTAGGG	RNL-7*	CAGCTATGTCCACCGGCTCA
RNL-124	GGAGATTCTGCCTCTGTTTG	RNL-78**	AAGCAACGGCCTCTAAGACA
RNL-125	GCAGGAAGACTCCTGAAT	RNL-136	GCCGATGAGTAACTCTTCTG
RNL-16*	ACCTGGAGATAGCTGGTCTT	RNL-137	GAGTTACTCATCGGCACTCC
RNL-126	GTAGGTGGGAGCCTACTGAT	RNL-138	GCTGTAGTTCCTGCTAAG
RNL-127	TCAGACCACTGAGCTGTTAC	RNL-139	GGTCATCTTGCCGAGTTCCT
RNL-128	CCACACGGTACTACGTACCT	RNL-140	TGAAGGAACTCGGCAAGATG
RNL-129	AGACGGTACGGCACTGGAAG	RNL-31**	TTMGTGCCGCCACTTATTAG
RNL-130	TGCCGTACCGTCTAGTAA	RNL-141	CTTATTAGTGGCACCCCTTC
RNL-131	AGCCTTTGGTTCGAGTCT	RNL-30**	TAGCATCGGGCAGGTATCAG
RNL-132	CACACGCCGAAACTAAAC	RNL-5*	GAGCTTCCTTTGCCATCCTA

Suppl. Table 2: Numbers of spores and clones analyzed to obtain partial or complete sequences from DNA extract of a single spore (* or of several spores) by directly sequencing or by cloning. When two approaches were used, samples directly sequenced were from approach 3.

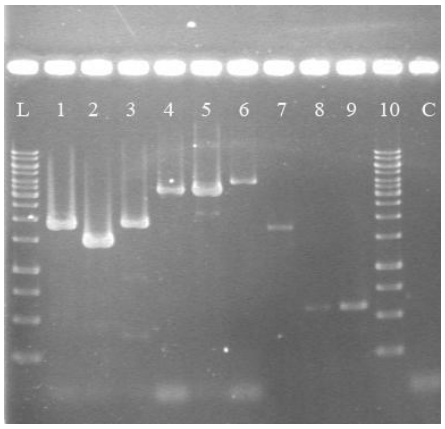
AMF	Mt-LSU			Nuclear rRNA
	Approach	Direct sequencing	Cloning	
Glomus coronatum (ZTL)	2	3 spores	1 spore (3 clones)	2 spores (1 clone each)
Glomus mosseae (ISCB18)	2	3 spores	2 spores	2 spores (1 clone each)
Glomus mosseae (ISCB13)	2	2 spores	2 spores	2 spores (1 clone each)
Glomus mosseae (ISCB14)	2	4 spores	3 spores	2 spores (1 clone each)
Glomus mosseae (BEG12)	2	2 spores	1 spore	2 spores (1 clone each)
Glomus caledonium (BEG20)	2	2 spores	1 spore	2 spores (1 clone each)
Glomus geosporum (BEG18)	2-3	1 spore	1 spore	1 spore (2 clones)
Glomus clarum (BEG142)	2	-	1 spore	1 spore (2 clones)
Glomus clarum (MUCL46238)	2	-	1 spore	1 spore (2 clones)
Glomus sp. (ISCB34)	3	-	2 spores	2 spores (1 clone each)
Scutellospora verrucosa (MN186)	1-3	1 clone (DNA from hyphae)	1 spore (3 clones)	1 spore (2 clones)
**Scutellospora castanea* (BEG01)	1-3	1 clone	Spores (4 clones)	Spores (1 clone) 2 spores (1 clone each)

Suppl. Table 3: Functional domains were from *G. geosporum* (1) and *G. clarum* isolates MUCL46238 (2) and BEG142 (3); disrupted domains from *G. coronatum* (4), *G. caledonium* (5), and *G. intraradices* isolates JJ291 (6), JJ141 (7) and JJ183 (8). N.A. means not applicable. Average of all pairwise comparisons: dS/dN = 1.6217.

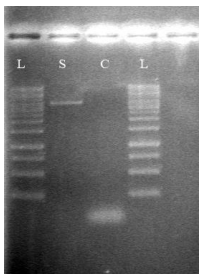
dS/dN value within functional domain	dS/dN value within disrupted domain	dS/dN value in *G. coronatum*
2.0119 (1-2)	1.2433 (4-5)	0.6994 (4-1)
2.0119 (1-3)	0.5547 (4-6)	0.7600 (4-2)
N.A. (2-3)	1.0238 (4-7)	0.8938 (4-3)
Mean value : 2.0119	0.6259 (4-8)	1.2433 (4-5)
	2.1563 (5-6)	0.5547 (4-6)
	2.4253 (5-7)	1.0238 (4-7)
	1.7540 (5-8)	0.6259 (4-8)
	1.0724 (6-7)	Mean value : 0.8287
	1.9714 (6-8)	
	1.0861 (7-8)	
	Mean value : 1.3913	



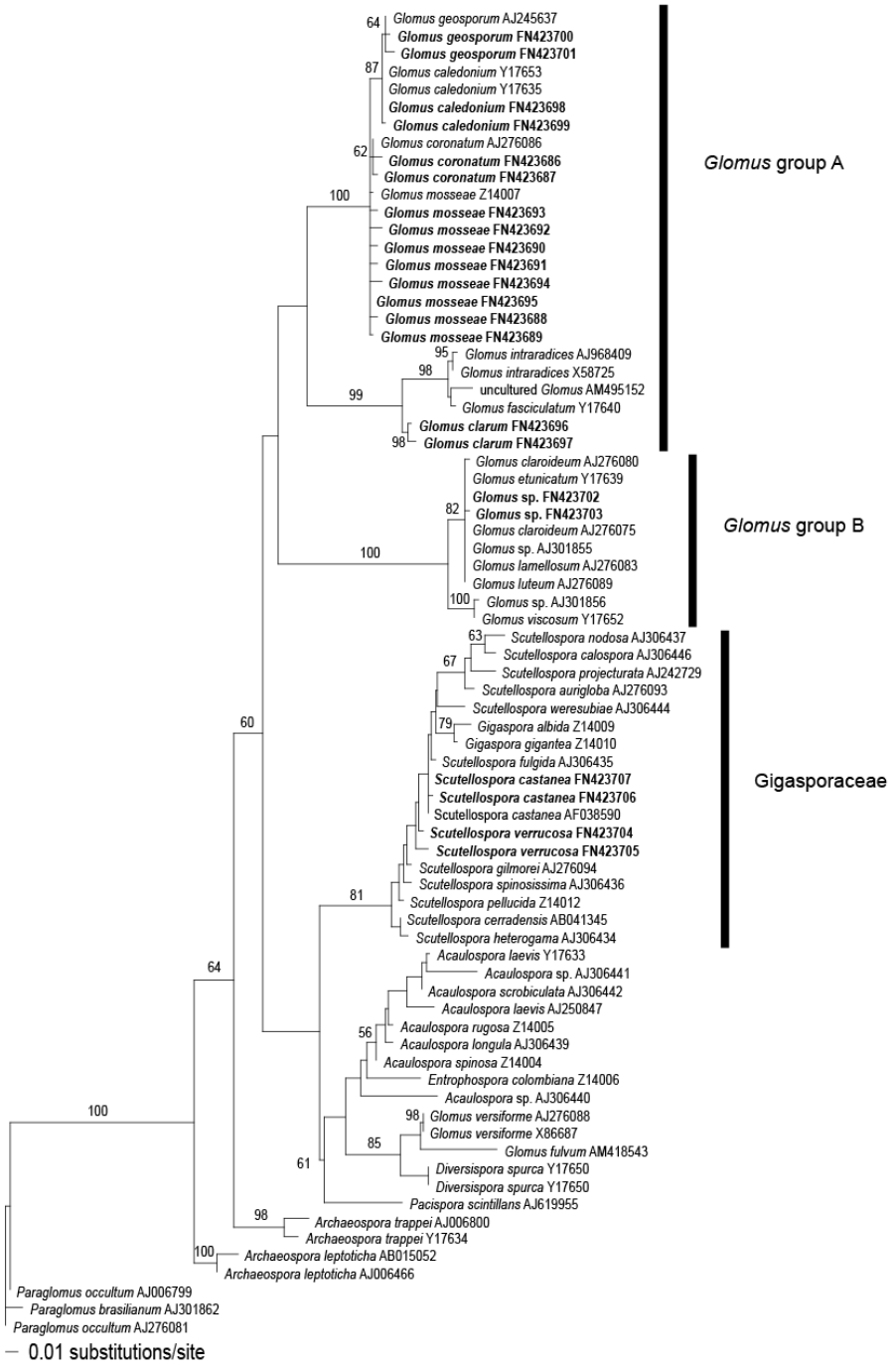
Suppl. Fig. 1: Location and orientation of additional RNL primers used in approches 1 and 2 and not required in approch 3. RNL126 was specific to Gigasporaceae and RNL-12/7/78/5 were only used to sequence *G. clarum*.



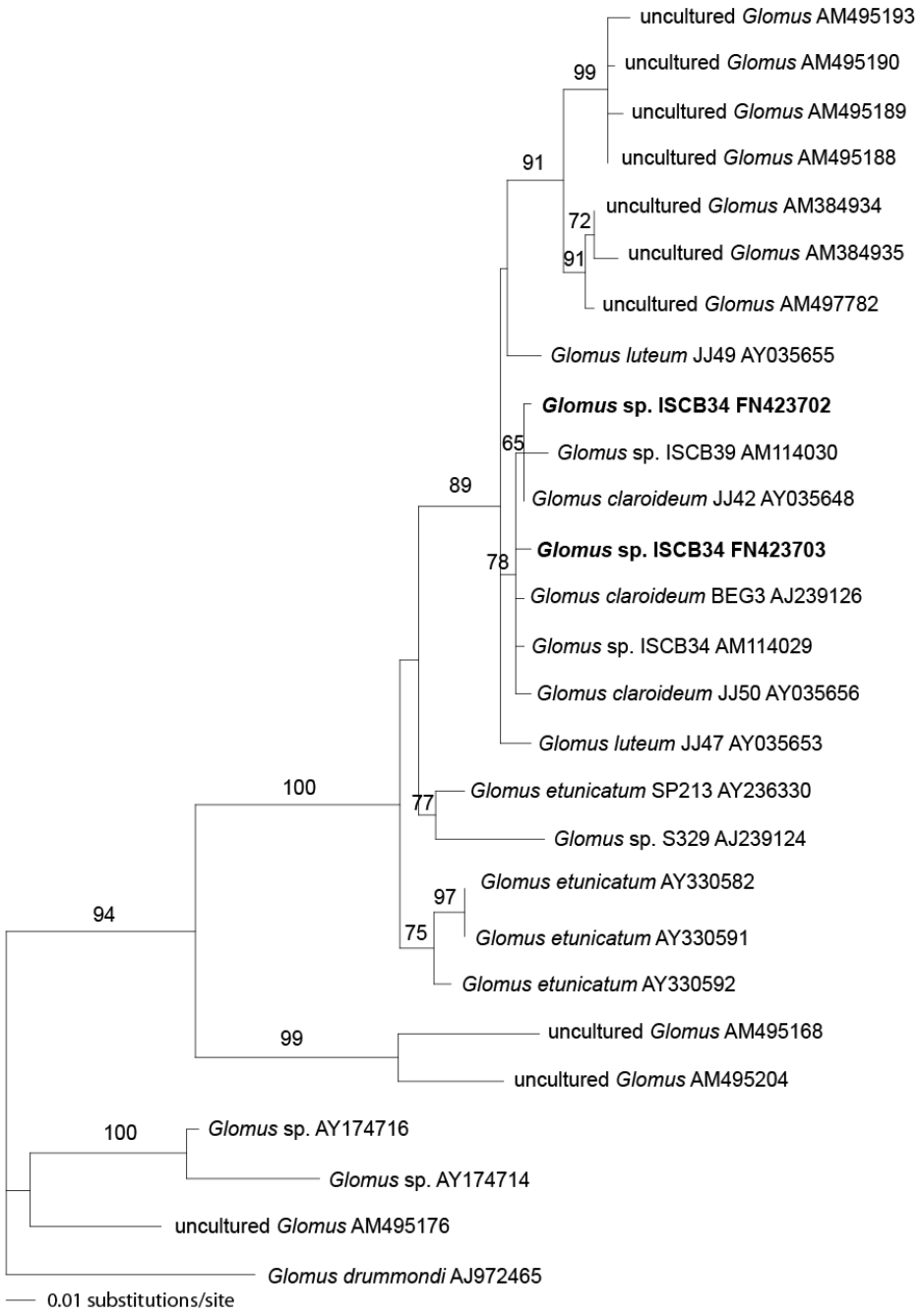
Suppl. Fig 2 : Agarose gel electrophoresis showing the length of the PCR products obtained using the primers 29/Ar5 for the following species : 1. *G. intraradices*, 2. *G. clarum*, 3. *G. geosporum*, 4. *G. mosseae*, 5. *G. caledonium*, 6. *G. coronatum*, 7. *G. etunicatum*-like, 8. *S. verrucosa*, 9. *S. castanea*. L indicates DNA ladder (bp): 250, 500, 750, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 5000, 6000, 8000, 10000. C shows the negative control.



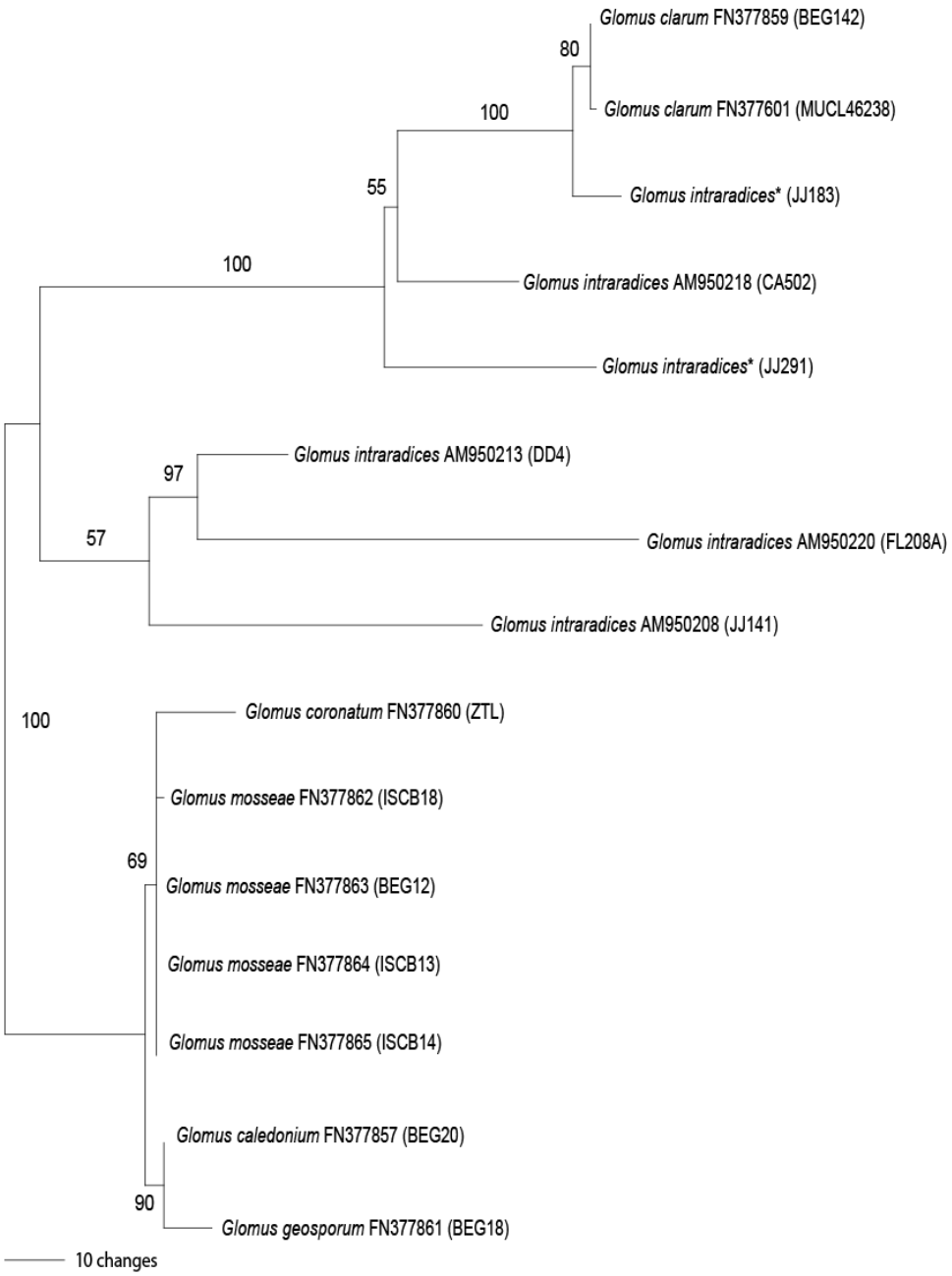
Suppl. Fig. 3: Agarose gel electrophoresis showing the PCR products of *Allium porrum* roots colonized by *G. coronatum* (S). C indicates the negative control and L DNA ladder (bp): 250, 500, 750, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 5000, 6000, 8000, 10000.



Suppl. Fig. 4: Phylogenetic tree of Glomeromycota based on 369 bp of 18S n-rDNA sequences. *Paraglomus brasilianum* was used as outgroup. The tree was generated by a heuristic search under the maximum likelihood criterion. Numbers on the nodes indicate neighbor-joining bootstrap values from 1000 replicates. Labels includespecies name, accession numbers and isolate codes (in brackets). Sequences obtained in the present study are shown in boldface.



Suppl. Fig. 5: Phylogenetic tree of *Glomus* group B based on 410 bp of 5.8S rRNA and ITS2 sequences. *Glomus drummondi* was used as outgroup. The tree was generated from a heuristic search under the maximum likelihood criterion. Numbers at the node indicate neighbor-joining bootstrap values from 1000 replicates. Labels include species name, accession numbers and isolate codes (in brackets). Sequences obtained in the present study are shown in boldface.



Suppl. Fig. 6: Phylogeny of Glomeromycota based on 1239 bp of intron 1149 sequences. Single most parsimonious obtained by a heuristic search under the parsimony criterion. Numbers on the nodes indicate bootstrap values from 1000 replicates. Labels include species name, accession numbers and isolate codes (in brackets). *, consensus sequences obtained from AM950210, AM950211 and AM950212 for *G. intraradices* JJ183; AJ973189, AJ973190 and AJ973191 for *G. intraradices* JJ291.

Chapter 3: Inter-isolate polymorphism of intergenic spacers in the mitochondrial genome of *Glomus intraradices:* potential for analyses of intraspecific genetic diversity

I. Abstract

In the ubiquitous group of obligate biotrophic arbuscular mycorrhizal fungi (AMF), the mitochondrial large subunit gene of ribosomal RNA (mtLSU) has been established as a molecular marker for population studies in *Glomus intraradices sensu lato* (corresponding to the nuclear ITS rDNA phylotype GLOM A-1). Some evidence suggested that mitochondrial genes are monomorphic with a fungal strain but vary among isolates. In this study, therefore, we targeted four mitochondrial intergenic spacers and assessed their polymorphism among six isolates of *G. intraradices* (*sensu lato*) that are known to differ in their mtLSU, a field sample of known mtLSU haplotype, and database sequences from two additional isolates. The highest resolution (six haplotypes) was achieved by the rns/nad5 spacer, a DNA fragment containing open reading frames (ORFs) of a putative GIY homing endonuclease (HE). Other spacer regions showed identical sequences among two isolates or more. In some cases, the respective regions could not be amplified as PCR parameters still need adjustments. Our data demonstrate that the intergenic spacers regions of the mitochondrial genome of AMF provide regions polymorphic between isolates and that their combined resolving power can be superior to mtLSU.

II. Introduction

Arbuscular mycorrhizal fungi (AMF; phylum: Glomeromycota) are crucially involved in plant nutrient absorption and plant community diversity and productivity (van der Heijden et al., 1998). Based on fossil evidence, the arbuscular mycorrhizal symbiosis is thought to be about 460 million years old (Redecker et al., 2000). Despite the well-established ecological relevance, little is known about the genetics of AMF. These fungi have been considered to be asexual since no reproductive structures have been observed (Gandolfi et al., 2003). The fungi produce resting spores which are multinucleate, but whether they are genetically identical (Pawlowska & Taylor, 2004) or different (Kuhn et al., 2001; Koch et al., 2004; Hijri & Sanders, 2005; Croll et al., 2009; Croll & Sanders, 2009) is still not yet clear.

The identification of fungal individuals, as well as the assessment of genetic diversity and gene flow are dependent on molecular genetic markers. The most commonly-used molecular markers in AMF research are rDNA sequences, but the existence of multiple variants complicates AMF species identification. In contrast to the heterogeneity of nuclear-encoded genes, the mitochondrial large subunit rDNA sequences (mtLSU) of AMF have been shown to lack intra-individual sequence polymorphism (Raab et al., 2005; Börstler et al., 2008; Börstler et al., 2010; Thiéry et al., 2010). MtLSU sequences were even revealed as powerful genetic markers for identification of 12 out of 16 isolates within *G. intraradices* species (Börstler et al., 2008). Molecular tools appropriate for population structure studies based on the nuclear genome were also made. In particular, microsatellites were analyzed for their high polymorphism level and 7 out of 8 isolates tested were distinguishable (Mathimaran et al., 2008a). Combination of microsatellites, nuclear gene intron and mitochondrial ribosomal gene intron markers allowed resolution of 18 genotypes out of 48 isolates (Croll et al., 2008b). There is an urgent need to develop tools for tracking isolates or individual clones of an AMF species, and despite the relevance of AMF diversity within species only few studies were published on genetic population structure.

Lee & Young (2009) who sequenced the complete mitochondrial genome laid the basis for developing further sensitive mitochondrial markers. Intergenic spacers are

assumed to be neutral molecular markers, affected to considerably less extent by selective pressure than protein-encoding genes. Intergenic regions exhibit polymorphism that is widely used for intraspecific and population studies e.g. in plants (Pleines et al., 2009), to observe the distribution of genetic variation within or among populations. Primers can be designed from protein-coding flanking genes that are relatively conserved and thus facilitate amplification of this non-coding DNA. For *G. intraradices*, Formey & Roux (2009) demonstrated that between the two isolates DAOM197198 and FACE494 there are some polymorphic regions.

The objective of the present study was to explore the polymorphism of selected intergenic regions of the mitochondrial genome to assess their suitability as molecular markers to distinguish isolates of *G. intraradices* (*sensu lato*).

III. Materials and methods

1. DNA extraction

This study used 7 DNA extracts already analyzed for the mitochondrial large subunit rRNA gene sequences for the study of Börstler et al. (2008; Table 1). The isolates JJ291, JJ183, and CC4 originated from root organ cultures (ROC) on transformed carrot roots (Bécard & Fortin, 1988). DNA of the isolates SW205, JA202, and KE114, were extracted from unpurified spore suspensions, containing additional organic matter. A DNA extract originated from roots of a field experiment in Frick (Switzerland), corresponding to the BEG75 mtLSU haplotype. The DNA extracts were either obtained (i) from 1-15 spores, following the procedure described in Redecker et al. (1997) or (ii) from > 15 spores, or 50-80 mg (wet weight) of organic matter (inocula/hyphae/roots), by using the DNeasy Plant Mini Kit (Qiagen, Hilden, Germany).

Table 1: Isolate code, geographic origin, and mitochondrial large subunit haplotype, according to (Börstler et al., 2008) of the studied isolates of *G. intraradices sensu lato*. * = data from Lee & Young (2009), accession number FJ648425; **= data from Grimwood (unpublished), accession number AC237367.

Isolate code - Origin	Haplotype
JJ291 - Switzerland	I
CC4 - Canada	II
root sample - Switzerland	III
SW205 - Switzerland	IV
JJ183 - Switzerland	V
JA202 - Japan	VIII
KE114 - Kenya	IX
FACE494 - Switzerland *	I
DAOM197198 - Canada**	I

2. PCR amplification of mitochondrial intergenic spacers

Four intergenic spacers were assessed for their suitability as molecular markers, the regions between the genes rns and nad5, atp6 and nad2, nad3 and nad6, as well as nad6 and cox3. For convenience, we hereafter refer to these spacers by their flanking genes in the format "spacer gene 1/gene 2". For instance, spacer rns/nad5 corresponds to the non-coding region in between the genes rns and nad5. Primers were newly designed based on the public flanking gene sequence information from isolate FACE494 of *G. intraradices* by using the software PrimerDesigner v.3.0 (Scientific & Educational Software, Cary, NC, USA). The PCR-amplified fragments contained the spacer regions between the genes mentioned above (Table 2; Fig. 2). PCR amplifications were carried out in a total volume of 50 µl with Phusion High-Fidelity DNA polymerase, 1 µl of genomic DNA, 1X Phusion HF buffer, 0.15 mM of dNTPs, 0.4 µM of each primer, 20 µg BSA and 3% DMSO on the Eppendorf Mastercycler epgradient S (Vaudaux-Eppendorf, Schönenbuch, Switzerland). Thermal cycling parameters were as follows: 30 s at 98°C, then 37 cycles of 10 s at 98°C, 20 s at 58°C, 100 s at 72°C and 10 min at 72°C.

Table 2: Newly designed primers for amplifying intergenic spacers, targeted to the 3' and 5' terminal ends of mitochondrial gene sequences. Numbers (1, 2, 3 or 4) preceding primer names designate the amplified regions between the genes rns and nad5, atp6 and nad2, nad3 and nad6, as well as nad6 and cox3, respectively. For spacer nad6/cox3 (region 4), "4.1" designates the region initially amplified and "4.2" the most polymorphic, diagnostic, region within the 4.1 region.

Primer name	**Sequence (5'-3')**	**Primer name**	**Sequence (5'-3')**
1.Rns1	GCGCACTAATTGTCCATCAG	4.1.Nad6-1f	ACTAGCGGTTCCACCTCTTA
1.Nad5-1	GAGGCCATAGAAGCTAGGAG	4.1.Cox3-1r	TAGAGGCCATGGTGAAGGTT
2.Atp6-1	GGAGGATCTCCTTCGACCTA	4.2.N6-1	AGGGCAAGGATGGAACTAAC
2.Nad2-1r	GCATGTTCCAAGCCAATACC	4.2.N6-2	GATAGCCAACCTTGGGAATC
3.Nad3-1	CCTGTAGGATTGGCCTATTG		
3.Nad6-1r	GGTAAGCTGAGACGGTAATG		

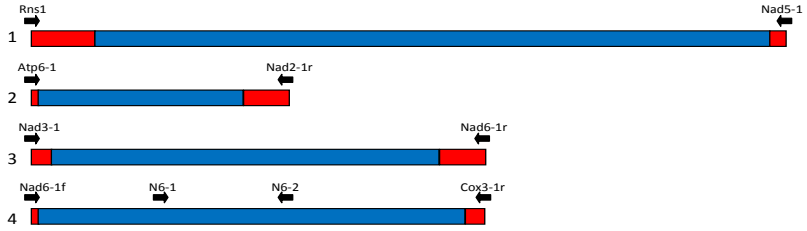


Fig. 2: Location and orientation of primers on schematic sequences of the region amplified for G. intraradices FACE494 between the genes rns and nad5 (1), atp6 and nad2 (2), nad3 and nad6 (3), and nad6 and cox3 (4). Coding regions are represented in red (▮), intergenic spacers in blue (▮). Diagram approximatively to scale.

3. Cloning and sequencing

Incubation at 72°C for 13 minutes with *Taq* polymerase, 2mM $MgCl_2$ and 0.125 mM dATP was used to adenylate the blunt-ended PCR products from the PCR amplifications with the Phusion polymerase prior to purification with the High Pure Kit (Hoffmann LaRoche, Basel, Switzerland). pGEM® Vector System I (Promega/Catalys, Wallisellen, Switzerland) was used following the manufacturer's instructions to clone the purified PCR amplicons for subsequent sequencing on ABI310, ABI3130xl and ABI3500 capillary sequencers (Applied Biosystems, Foster City, CA) with the vector primers M13f (GTAAAACGACGGCCAGTG) and M13r (GGAAACAGCTATGACCATG) and Big Dye Terminator v.3.1 chemistry.

4. Sequence analysis

All newly-generated sequences were edited and manually aligned to available sequences in Genbank (accession number FJ648425 and AC237367) using Bioedit (Hall, 1999).

The rates of synonymous (dS) versus non-synonymous (dN) substitutions were computed using the Synonymous Non-synonymous Analysis Program (Korber, 2000).

IV. Results

A region of each intergenic spacer selected was amplified from isolates JJ291, BEG75, SW205 and JJ183. However, some PCR primers designed from other regions did not give rise to amplicons or reliable sequence for some of the strains. This was, for instance, the case for isolate JA202, for which the PCR amplicon could not be aligned with other sequences for spacer rns/nad5. No atp6/nad2 sequence could be generated for the isolates KE114 and CC4, and no amplification of the spacer nad3/nad6 was obtained for isolate CC4. Only amplifications with the primers targeted to the spacer nad6/cox3 were successfully applied to the field sample and all the six tested isolates of *G. intraradices*. However, this spacer nad6/cox3 could only resolve isolates JA202, JJ291 and JJ183 from the field sample and the other two isolates. Spacer rns/nad5 discriminates

the field sample and all the tested isolates, except for JA202 for which further investigations are needed. Details about the nucleotide sequences for each intergenic region are given below.

Table 3: Lengths (in base pairs (bp)) of the sequences obtained for each intergenic spacer targeted for different isolates of *G. intraradices* used in this study, which do not always comprise the complete amplicon. * not alignable sequence; ** no sequence obtained; - no sequence available in Genbank. Lengths of the sequences from the dagnostic region "4.2" are shaded. In the composite haplotype, each of the 4 numbers corresponds to a haplotype of the four loci, the spacers rns/nad5, atp6/nad2, nad3/nad6 and nad6/cox3, respectively. The number in brackets remains to be confirmed.

Isolate/Haplotype number	rns/nad5	atp6/nad2	nad3/nad6	nad6/cox3	Composite haplotype
JJ291/I	2092	690	1240	1203	1111
JJ183/V	1965	700	1188	1239	2222
SW205/IV	2195	690	1240	1189	3113
CC4/II	1153	**	**	246	4--4
JA202/VIII	*	575	1202	320	-3(4)5
KE114/IX	1020	**	1976	246	5-34
Root sample/III	1148	700	1201	246	6244
FACE494/I	2092	690	1240	1239	1112
DAOM197198/I	2092	-	-	-	1---

Spacer rns/nad5:

Fig. 3: Alignment of two variable non-coding spacer regions between the genes rns and nad5 potentially useful for discriminating strains of *G. intraradices*. Dots indicate identical nucleotide states and shaded hyphens stand for gaps. *, RS refers to root sample.

The nucleotide sequence of spacer rns/nad5 revealed differences among the root sample and all the isolates JJ291, JJ183, SW205, KE114, and CC4. The amplified region showed deletions ranging from 1047 to 1062 bp in the isolates CC4, KE114 and the root sample, compared with the isolates JJ291, JJ183 and SW205 (Table 3). The isolate CC4

and the root sample differed only by substitutions of a few nucleotides. Isolate KE114 displayed substantial nucleotide sequence variation in the spacer rns/nad5, besides deletions, when compared to the root sample and all other isolates. A 39-bp insertion and 34-bp and 131-bp deletions were identified in isolate JJ183 in comparison to the isolates JJ291 and SW205. Isolate SW205 also differed to other isolates in having 61-bp and 42-bp insertions. No sequence differences were observed between isolates FACE494 and DAOM197198 and the isolates JJ291 and FACE494 differed in four single nucleotide polymorphisms. Despite minor sequence variation in the intergenic spacer region, these two isolates show two open reading frames (ORFs) of 193 amino-acids (aa) and 122 aa for GIY-YIG endonucleases. Isolate SW205 also shows an ORF of 193 aa for the first, but the second ORF of the endonuclease is disrupted by an 42-bp insertion in its conserved domain. Similarly, ORFs of the GIY-YIG endonuclease were found for the isolates CC4 (184 aa) and the root sample (180 aa) as well as ORFs in which the putative conserved domain is missing. Isolate KE114 displays substantial sequence variation and possesses an ORF of 147 aa for the GIY-YIG endonuclease. In isolate JJ183 the putative protein is only 89 aa long, since, compared with the isolates JJ291 and FACE494, there is only a part of the second ORF of the endonuclease present.

The overall ratio of the rates of synonymous to non-synonymous substitutions, dS/dN, was higher than 1, which suggests that negative selection maintains the functionality of the endonuclease in favor of the "homing" event, which is the horizontal transfer of the intron-containing allele encoding endonucleases to cognate alleles that lack that element.

Spacer atp6/nad2:

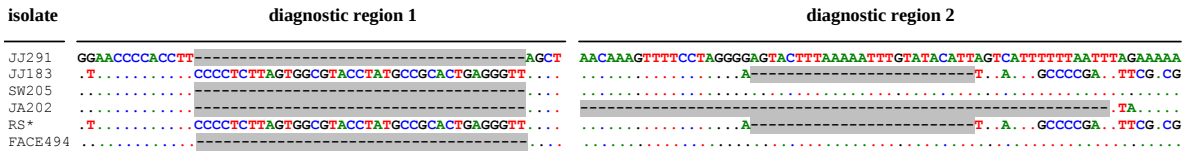


Fig. 4: Alignment of two variable sequence regions of the intergenic spacer atp6/nad2 potentially useful for discriminating isolates of *G. intraradices*. Dots indicate identical nucleotide states and shaded hyphens denote gaps. *, RS refers to root sample.

A 115-bp deletion in the intergenic spacer atp6/nad2 distinguishes the isolate JA202 of *G. intraradices* from all other tested isolates. No single nucleotide polymorphism could be found in this region for the isolates JJ291, SW205, and FACE494 and for isolate JJ183 and the root sample, although the latter two show a 37-bp insertion and 25-bp deletion, compared to the former set of isolates.

Spacer nad3/nad6:

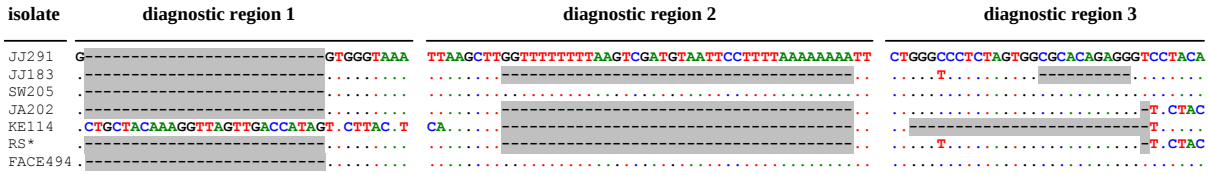


Fig. 5: Alignment of two variable regions of the intergenic spacer nad3/nad6 potentially useful for discriminating isolates of *G. intraradices*. Dots indicate identical nucleotide states and shaded hyphens stand for gaps. *, RS refers to root sample.

Isolate KE114 was characterized by a 26-bp insertion followed by a highly polymorphic region of 120 bp. A large insertion of 801 bp explains the uniquely large size of 1976 bp of the amplicon of isolate KE114 (Table 3), despite the presence of 26-bp and 29-bp deletions. Isolate JJ183 differed from other isolates only by short indels, including a 10-bp deletion. Isolates JJ291, SW205 and FACE494 can be distinguished from other isolates by a 38-bp insertion. Isolate JA202 and the root sample displayed minor sequence differences among each other of 3 bp.

Spacer nad6/cox3:

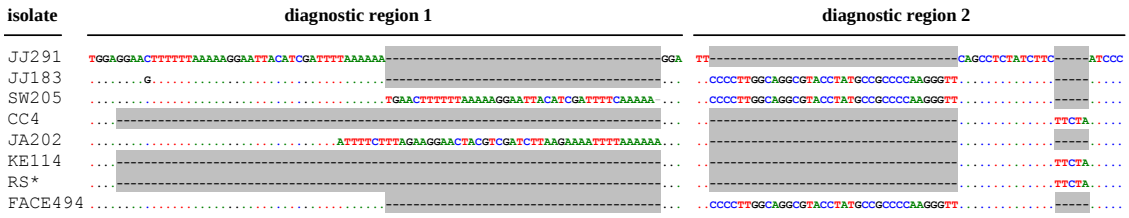


Fig. 6: Alignment of two variable intergenic spacer regions of the genes nad6 and cox3. Dots indicate identical nucleotide states and shaded hyphens represent gaps. *, RS refers to root sample.

Initially-used primers for the spacer nad6/cox3 were first applied to the three isolates JJ291, JJ183, and SW205 and the sequences obtained from these products were then used to design new primers, targeting the most polymorphic region. This variable region was then also sequenced from the additional isolates CC4, JA202, KE114, and BEG75. The region of the intergenic spacer with the highest polymorphism consisted of two indel-containing regions, the variation of which could be used to discriminate between the isolates JJ291, SW205, FACE494 and JA202 (Fig. 6). The spacer nad6/cox3 was, however, invariable among the isolates CC4, KE114, and the root sample and between isolates FACE494 and JJ183. More precisely, the isolates JJ291, JJ183, FACE494, SW205, and JA202 contained insertions of 39 bp, 39 bp, 39 bp, 78 bp, and 79 bp, respectively, in the first distinguishing region, which were missing in the isolates CC4, KE114, and the root sample. The isolates JJ183, FACE494, and SW205 contained a 36-bp insertion, whereas the isolates CC4, KE114, and the root sample possessed a 5-bp insertion relative to the other isolates in the second variable region (Fig. 6).

V. Discussion

Four intergenic spacers of mitochondrial genes were explored in six isolates and a root sample of known mtLSU haplotype of *G. intraradices* with regard to their suitability to serve as markers in future population studies. Their partial non-coding nature and thus relaxed selective constraints were expected to make them potential high-resolution molecular markers. Indeed, the analyzed set of fungal isolates showed some sequence polymorphism both by indels as by point mutations.

None of the individual intergenic spacer regions studied here resolved all of the analyzed isolates of *G. intraradices*. However, in combination they allowed to distinguish mtLSU haplotypes and even isolates corresponding to the same mtLSU haplotype. Indeed, the isolates JJ291, FACE494 and DAOM197198 from Canada, currently being used for genome sequencing (Martin et al., 2008) belong all to the widespread mtLSU haplotype I and distinction between JJ291 and FACE494 was made possible by the combination of the intergenic spacer (Fig. 7). Microsatellites used by Croll et al. (2008b) and Mathimaran et al. (2008a) revealed a higher genetic diversity than the mtLSU marker system, displaying genetic variation between JJ291 and DAOM197198. However, sequence data for isolate DAOM197198 was limited to spacer rns/nad5 and therefore not considered in our analysis. The highest level of information was provided by the spacer rns/nad5, distinguishing the field sample and all five isolates amplified except for JA202 and thus should receive most attention in attempts to develop new marker systems. Fig. 7 shows that in many cases, two markers were sufficient to differentiate the six locally available isolates analyzed by PCR and the haplotype from the colonized root sample. Nevertheless, the results show that the ribosomal mtLSU gene fragment used by Börstler et al. (2008) is in fact well-suited to distinguish individual isolates of *G. intraradices*, because, the six isolates analyzed by PCR were distinguishable by the mtLSU gene sequence, which is something that was not possible with any of the four investigated intergenic spacers alone.

These are promising, preliminary results that need to be further confirmed, especially for some of the strains in which no sequences could be generated with the employed PCR primers, or with polymorphic regions so long that they were not included in the alignment. Problematic amplifications might be attributed to base pair changes in the annealing sites, which could be circumvented by modifying or shifting the priming sites, even tough adjustements of PCR parameters might solve the problems. However, changes in gene order cannot be completely ruled out as nothing is currently known about the conservation of this feature apart from the two, relatively closely related strains FACE494 and DAOM197198 (Formey & Roux, 2009).

The fact that the correct amplicons were obtained from DNA from roots of a field experiment, suggests that the designed primers have a high specificity and can be applied

not only to pure DNA extracts of AMF but also to environmental samples, opening up the possibility to assess the genetic diversity of *G. intraradices* in the field. Moreover, the amplified region did not show any similarity with other sequences from the BLAST network server (NCBI; www.ncbi.nlm.nih.gov). However, as with all multilocus markers, linking the gene markers to obtain an organismal genotype is problematic if several genotypes are present in the same sample.

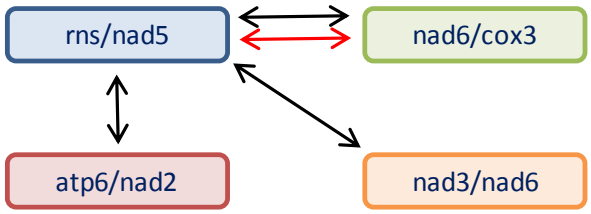


Fig. 7: Diagram showing which two spacers should be combined to resolve (i) ⟷ all of the mtLSU haplotypes, (ii) ⟷ all of the 7 analyzed isolates and the field sample.

Isolate KE114 was unique among *G. intraradices* isolates in the study of Börstler et al. (2008) because it was the only isolate not containing any of the three mtLSU introns out of the 16 isolates of *G. intraradices* analyzed. This absence of molecular markers complicated the determination of its genetic relationships. Interestingly, isolate KE114 in the present study was the isolate showing most unique sequence features, suggesting it is genetically quite distant from the others.

In comparison to nuclear-encoded molecular markers, mitochondrial markers in AMF provide the advantage that their homogeneity within the organism is well-established and that they originate from known gene regions. This facilitates the verification of the origin of such markers and the design of primers from other species. In contrast, some microsatellite markers originated from genomic regions randomly sequenced (Mathimaran et al., 2008b) and their mutation rates and therefore their suitability can vary depending on their location in coding or non-coding regions.

The spacers analyzed in this study were chosen according to their length and their position between conserved genes. The mitochondrial genome offers several additional

regions that will be useful to explore for marker development in order to refine mitochondrial genotyping in *G. intraradices*. The spacer rns/nad5 that contains putative ORFs of a GIY endonuclease in all the studied isolates displays considerably more polymorphism between isolates, compared to the three others spacers. This finding of elevated sequence polymorphism around another genomic region encoding an endonuclease is in good agreement with our previous finding of higher levels of sequence polymorphism, associated with the horizontally transferred endonuclease in one of the introns of the mtLSU (Thiéry et al., 2010). Therefore, we suggest that regions harboring endonucleases may be promising candidates to develop marker systems for distinguishing isolates of *G. intraradices*. Clearly, rapidly evolving genomic regions, such as those associated with genes of endonucleases, should receive attention in the development of highly-resolving population-level genotyping approaches.

Chapter 4: Mitochondrial large subunit rDNA haplotype diversity of *Glomus intraradices* in roots from geothermal and non-geothermal soils of Yellowstone National Park and Iceland

I. Abstract

Arbuscular mycorrhizal fungal (AMF) populations colonizing roots in geothermal and non-geothermal sites were investigated under the hypothesis that they may confer adaptive advantages to their host plants under extreme environmental conditions. Research was focused on sequence identity and distribution of mtLSU haplotypes of *Glomus intraradices sensu lato* (corresponding to the nuclear ITS rDNA phylotype GLOM A-1), since this was the most frequently sampled taxon. The study areas for root sampling were geothermal and non-geothermal soils at Yellowstone National Park (YNP, USA) and in Iceland. Nine different restriction fragment length polymorphisms were recorded in 44 root samples of three plant species at both geothermal and ambient sites, corresponding to 15 sequence haplotypes. Eighty percent of the mtLSU haplotypes had never been found previously. Intra type and sequence haplotype richness were higher at YNP, compared to Iceland. Most noticeable, however, no overlap in the distribution of haplotypes was found between the two geographical locations. Haplotype I, commonly found in arable fields in Switzerland, was recorded at thermal sites in Iceland, which may be related to its ruderality and adaptation to physical disturbance. Interestingly, additional closely related haplotypes of haplotype I were only found in Iceland, but not at YNP. These results suggest, firstly, that extreme thermal habitats may be inhabited by strains of *G. intraradices*, which also occur in arable fields, and secondly, that additionally there may also be strong biogeographical structuring in the occurrence of *G. intraradices* strains.

II. Introduction

Arbuscular mycorrhiza (AM) is a ubiquitous symbiosis formed between most land plants and fungi from the phylum Glomeromycota (Schüßler et al., 2001a). All members of Glomeromycota are obligate symbionts that improve the mineral nutrition of their carbohydrate-providing partners. Strains or higher taxa of AM fungi (AMF) have been shown to differ in tolerance to acid soil conditions and high Al^{3+} concentrations (Clark, 1997). Soils in geothermal areas are commonly characterized by having elevated temperatures and high concentrations of manganese, iron, aluminium, arsenic, and sulfide (Henley, 1986; Rodman, 1996), often impairing plant growth. Nonetheless, many plants have evolved strategies to cope with all these stresses, such as an altered root morphology, expression of heat-shock proteins (Stout et al., 1997), use of water and air humidity (Germino & Wraith, 2003), growth within moss mats (Tercek & Whitbeck, 2004), and formation of AM (Bunn & Zabinski, 2003; Bunn *et al.*, 2009). Furthermore, a virus infecting the endophytic ascomycete, *Curvularia protuberata*, was shown to confer heat tolerance to the tropical panic grass *Dichanthelium lanuginosum* (Marquez et al., 2007) in Yellowstone National Park (YNP), USA. YNP is a geologically very old geothermal area, characterized by volcanic activity, geysers, fumaroles, hot springs, and mud pods.

Plants in geothermal areas across YNP were reported to be colonized by AMF, despite acidic soils and elevated temperatures in the rooting zone, based on microscopic investigations (Bunn & Zabinski, 2003) and molecular genetic analyses (Appoloni et al., 2008). Appoloni et al. (2008) used nuclear-encoded rRNA gene markers to characterize the structure and composition of AM fungal assemblages in roots of the grasses *Dichantelium lanuginosum* and *Agrostis scabra* from thermal and non-thermal areas in YNP and roots of *Agrostis stolonifera* from Iceland. Composition of the AMF communities differed among sites and host plant species. The highest number of phylotypes was reported for the lineage *Glomus* group A, with *Glomus intraradices* GLOM A-1 being the most frequent phylotype. The definition of *G. intraradices* is ambiguous (Stockinger et al., 2009), and we refer to "*G. intraradices*" for "*G. intraradices* GLOM A-1", a phylogenetic lineage frequently detected in environmental studies (Hijri et al., 2006; Sýkorová et al., 2007b). The phylotaxon *G. intraradices* GLOM A-1 also contains the fungal strain DAOM197198 whose genome is currently being sequenced (Martin et al., 2008), but not the type strain (FL208) from Florida (Schenck & Smith, 1982).

G. intraradices GLOM A-1 appears to be the most common and widespread phylotaxon with a distribution across a broad range of habitats. It was detected for instance in mountain and alpine meadows (Börstler et al., 2006; Sýkorová et al., 2007b), calcareous grasslands (Sýkorová et al., 2007a), phosphate-polluted sites (Renker et al., 2005), arable fields (Hijri et al., 2006), and, indeed, at geothermal sites (Appoloni et al., 2008). Concluding from its distribution, the organisms giving rise to the sequences, corresponding to phylotype *G. intraradices* GLOM A-1, must exhibit a generalist life history strategy. Owing to the fact that many fungi belonging to *G. intraradices* GLOM A-1 can be cultured and sporulate fast and abundantly, this phylotaxon is often experimentally used as an AMF model, also for functional analyses.

In contrast to the homogeneity of mitochondrial genomes within individual strains of AMF (Raab *et al.*, 2005; Lee & Young, 2009; Thiéry *et al.*, 2010), there appears to be substantial variation among strains and species (Raab *et al.*, 2005; Börstler *et al.*, 2008). This has been shown for the mitochondrial large subunit rDNA sequences (mtLSU) of *G. intraradices* by Raab et al. (2005) and Börstler et al. (2008). Lack of intra-individual sequence variation in the mitochondrial genomes contrasts considerably with pronounced nuclear rDNA sequence variation in AMF (Sanders *et al.*, 1995; Lloyd MacGilp *et al.*, 1996). The suitability of mtLSU markers for field studies on *G. intraradices* in roots has recently been shown (Börstler et al., 2010).

The aim of the present study was to investigate intraspecific genetic diversity of the phylotaxon *Glomus intraradices* GLOM A-1 in geothermal areas of YNP and Iceland using the mtLSU PCR-RFLP-sequencing approach developed by Börstler et al. (2008). The research questions specifically addressed were:

- Do the extreme chemical and physical conditions of geothermal sites decrease sequence haplotype richness of *G. intraradices*?
- Do geothermal sites harbor unique strains (mtLSU sequence haplotypes) of *G. intraradices*? Is there geographical structuring or possibly isolation by distance between the geothermal areas of YNP and Iceland?

III. Materials and methods

1. Field sites and root sampling

Yellowstone National Park, USA: Samples were collected from three geothermal areas and their surrounding zones at YNP by Appoloni et al. (2008): Rabbit Creek (Midway Geyser

Basin, 44°31' N 110°49' W), LG1, and LG2 (Upper Geyser Basin in the Lone Star Geyser Area, 44°34' N 110°48' W). The two collection basins were about 10 km apart. Roots of *Dichantelium lanuginosum* were sampled in geothermal areas, whereas such of *Agrostis scabra* could be collected at both geothermal and non-geothermal zones. At LG1 and LG2, the non-thermal control plants (Poaceae) were sampled outside the thermally-influenced areas. Further details about the sites can be found in Bunn & Zabinski (2003) and Appoloni et al. (2008). At the Rabbit Creek site, the upper soil temperature ranged between 31 and 42°C and soil pH between 3.4 and 4.8 in water. Geothermal soils contained four times more iron, but much less zinc and manganese than the transition area soils (Bunn & Zabinski, 2003). At LG1 and LG2, soil temperature reached even 50°C, and the soil pH was about 4.

Iceland: Roots of *A. stolonifera* were collected from two geothermal areas in Iceland (Appoloni et al., 2008): Reykjavik, Gunnuhver (Reykjanes Peninsula, 63°49' N 22°41' W) and Ölkelduháls (north of Hveragerdi, 64°03' N 21°14' W). At Reykjavik, soil pH was about 5 and soil temperature ranged from 38-55°C in thermal and 25-30°C in non-thermal soils. At Ölkelduháls, soil pH ranged from 4-4.5 and soil temperature from 37-41°C in thermal and from 24-28°C in non-thermal soils. The distance between the two sites was approximately 75 km.

Roots of the collected plants were washed, aliquoted (up to 100 mg/tube) and frozen at -80°C, some roots from YNP were lyophilized instead.

2. DNA extraction

DNA was extracted from root samples using a DNeasy Plant Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. DNA was eluted in a final volume of 100 µl. The 20 DNA extracts of roots from YNP used in this study had already been analyzed with regard to AMF community composition by Appoloni et al. (2008), as well as three out of a total of 24 DNA extracts from roots from Iceland.

3. PCR amplification of mtLSU marker fragments

MtLSU was amplified, using the nested PCR approach developed by Börstler et al. (2008, 2010). The primer pairs RNL-28a/RNL-5 and RNL-29/ RNL-30 were used for the first and second step of the nested PCR, respectively. The first PCR amplicons were obtained with 0.02 U/µl Phusion High-Fidelity DNA polymerase (Finnzymes, Bioconcept, Allschwil,

Switzerland) from about 1 µl of genomic DNA, using 1X Phusion HF buffer, 0.2 mM of dNTPs, 0.5 µM of each primer, 10 µg BSA, 3% DMSO. After initial denaturation for 30 s at 98°C, thermal cycling followed the program 10 s at 98°C, 20 s at 57°C, 1 min 40 s at 72°C for 37 cycles and 10 min at 72°C. PCR products of this first amplification step were diluted 1:100 in water and used as template in the second nested reaction. The second amplification round was carried out in a total volume of 50 µl with again 0.02 U/µl Phusion High-Fidelity DNA polymerase, 1 µl of 1:100 diluted PCR amplicon, 1X Phusion HF buffer, 0.15 mM of dNTPs, 0.4 µM of each primer, 20 µg BSA, and 3% DMSO. Thermal cycling started with 30 s at 98°C and proceeded through 36 cycles of 10 s at 98°C, 20 s at 60°C, 1 min 40 s at 72°C and 10 min at 72°C.

4. RFLP analyses

RFLP analyses followed the procedures outlined in Börstler et al. (2008, 2010). Briefly, each mtLSU PCR amplicon from DNA templates of environmental root samples was separately digested with *Dra*III, *Bsa*JI, and *Hin*dIII and loaded in this order onto the agarose gel. The obtained RFLP patterns that corresponded with *G. intraradices* GLOM A-1 were categorized as so-called "Intra types", following Börstler et al. (2010). Some mtLSU sequence haplotypes give rise to the same Intra type, meaning that RFLP analyses do not always provide full resolution for all mtLSU sequence variants. More details on definition of mtLSU RFLP patterns can be found in Börstler et al. (2010).

5. Cloning, sequencing, and analyses

An adenine overhang was added to the 3' blunt end of PCR products generated by DNA strand synthesis with Phusion polymerase, using a 13 minute incubation at 72°C with *Taq* polymerase, 2mM $MgCl_2$ and 0.125 mM dATP. Purified PCR products (High Pure Kit, Hoffmann LaRoche, Basel, Switzerland) were cloned into the pGEM® Vector System I, according to the manufacturer's instructions (Promega/Catalys, Wallisellen, Switzerland). Cloning was necessary for sequence analysis in cases of samples, which gave rise to mixed RFLP patterns, and simultaneously assisted sequencing, because sufficient template could be generated. PCR amplicons of vector inserts (primers M13f: 5'-GTAAAACGACGGCCAGTG-3', M13r: 5'-GGAAACAGCTATGACCATG-3') were cycle

sequenced after purification (Table 1), using Big Dye Terminator v.3.1 chemistry on ABI310, ABI3500, or 3500xL capillary sequencers (Applied Biosystems, Foster City, CA).

The Restriction Enzyme Mapping Application software (http://bioperl.macaulay.ac.uk/) was used for *in silico* predictions of the restriction fragment sizes (Table 2) and composite, predicted RFLP / Intra type patterns were drawn, using MS Excel (Fig. 1).

The sampling effort curve was calculated using the program EstimateS 8.0.0 (Colwell, 2005), sample order was randomized in 100 replications.

6. Phylogenetic sequence analyses

All newly generated mtLSU sequences were manually aligned to pre-existing multiple sequence alignments (Raab et al., 2005; Börstler et al., 2008; Börstler et al., 2010) in Bioedit (Hall, 1999). Phylogenetic trees were generated from distance, parsimony, and maximum likelihood analyses, employing the software PAUP*4b10 (Swofford, 2001).

Table 1: Newly designed sequencing primers for mtLSU PCR amplicons. Other employed sequencing primers were published by Raab et al. (2005) and Börstler et al. (2008, 2010).

Primer	Sequence (5'-3')
RNL-143	GCACGTATAGACGAGCATTG
RNL-144	TATATGCTCCGGCGTAGAGT
RNL-146	TGCTCCTGCGTGTATAACTG
RNL-147	TGCCGAGTTCCTTCAAGACG
RNL-148	ACAAGGCAGACTACGGCTAT
RNL-149	ACCGACACACAGGTCTGCAAGT
RNL-150	TGTTCTTCTGGGCCTTGC
RNL-151	GCCGACTGATCACACGCAAT
RNL-152	GGTCATCTTGCCGAGTTCCT
RNL-153	CATAGGTACGCCTAAGAG

Table 2: Restriction fragment sizes calculated *in silico* by simulated digestions of a region of the mitochondrial rRNA large subunit gene (mtLSU) amplified by the PCR primer pair RNL-29/RNL-30 with the three restriction endonucleases *Dra*III, *Bsa*JI, and *Hin*dIII.

Samples			**Complete**	***Dra*III**	***Bsa*JI**	***Hin*dIII**
Accession number	**Haplo-Type***	**Intra-Type***	**fragment length** (bp)	5'...CACNNN▼GTG...3'	5'...C▼CNNGG...3'	5'...A▼AGCTT...3'
S31/S32	I	1	2569	1430, 1139	1289, 1175, 105	2067, 502
S13-7	XIX	3a	1500	1127, 373	829, 448, 223	998, 502
S3	XXIV	24	2259	1886, 373	2036, 223	948, 585, 508, 218
S1	XXXVI	36	2912	1475, 1437	1416, 1181, 209, 106	1657, 751, 308, 196
S4/S6	XXXVII	37	2988	1661, 1327	1072, 905, 697, 209, 105	1548, 748, 496, 196
S5	XXXVIII	38	3192	1743, 1449	1416, 1193, 477, 106	1938, 751, 307, 196
S12	XXXIX	39	2973	2088, 885	2238, 443, 185, 107	1121, 885, 508, 459
S7/S33/S37/S47/S8	XL	40	2607	1468, 1139	1289, 1213, 105	2105, 502
S34/S35/S36/S49/S39/S43	XLI	40	2645	1506, 1139	1289, 1251, 105	2143, 502
S41	XLII	40	2646	1507, 1139	1289, 1252, 105	2144, 502
S44/S46	XLIII	38	3184	1752, 1432	1417, 1176, 485, 106	1929, 751, 308, 196
S13-8	XLIV	5c	2567	1449, 1118	1193, 829, 439, 106	2065, 502
S50	XLV	40	2645	1506, 1139	1289, 1250, 106	2143, 502
S2-4	XLVI	36	2919	1475, 1444	1416, 1188, 209, 106	1664, 751, 308, 196

Note: Haplotypes denote different sequence types and Intra types different RFLP patterns.

IV. Results

1. Detection of mtLSU haplotypes

MtLSU PCR amplicons of *G. intraradices sensu lato* (= nuclear ITS rDNA phylotype GLOM A-1) could be generated from 10 of the 15 genomic DNA extracts (66%) of mycorrhizal roots, previously analyzed for AM fungal community composition and found to yield PCR products of GLOM A-1 by Appoloni et al. (2008). The lengths of the mtLSU amplicons ranged between 1500 and 3184 bp, mainly depending on intron lengths (Table 3). Five of the ten root samples from non-thermal areas at YNP yielded mtLSU amplicons of *G. intraradices*. Two samples gave rise to either two sequence haplotypes affiliated with *G. intraradices* or one with another AM fungus. Two samples reported problematic for reproducible amplification of nuclear rDNA of AMF were twice tested negative in amplification trials for the mtLSU marker. Only three out of 10 samples from thermal areas at YNP yielded mtLSU PCR amplicons. The two samples in which sequences affiliated with the nuclear ITS clade GLOM A-1 had been found by Appoloni et al. (2008) yielded also amplicons for the mtLSU PCR-RFLP assay. Ten and eight of the 12 root samples yielded mtLSU amplicons for thermal and non-thermal areas in Iceland, respectively. That means that regardless of what would have been found with the nuclear ITS marker system, 75% of the newly prepared DNA extracts from Iceland yielded an mtLSU PCR amplicon.

Table 3: Sequence structure of the mitochondrial rRNA large subunit gene (mtLSU) of *G. intraradices* obtained using the PCR primer pair RNL-29/RNL-30. The RFLP patterns defined as "Intra types" and sequence haplotypes obtained from the analysis of 44 root samples are listed. Primer sequences were excluded for fragment size calculations. Introns containing putative open reading frames of LAGLIDADG homing endonucleases are shaded. All analyzed samples from YNP originated from Rabbit Creek (RC) except for two from either LG1 or LG2. The samples from Iceland were collected at Ölkelduhals or Reykjanes (indicated in brackets). Host plant species are indicated in brackets, following the sample identifiers: *Dichantelium lanuginosum* (Dl), *Agrostis scabra* (Asc), *Agrostis stolonifera* (Ast), Poaceae (P). *, mixed patterns.

Sample origin, identifier and host plant identity	**Clones (accession number)**	**Intra type**	**Haplotype**	**Introns**			**Exons**	**Amplicon length** (bp)
				Pos. 1 (length, bp)	Pos. 2 (length, bp)	Pos. 3 (length, bp)	Complete/parts between intron positions (bp)	
YNP - thermal soil								
3 (RC - Dl)		24	XXIV	No	339	809	1071/363,670,38	2219
6 (RC - Dl)		37	XXXVII	961	No	942	1046/357,651,38	2949
12 (RC - Asc)		39	XXXIX	512	No	1350	1071/363,670,38	2933
YNP - non-thermal soil								
1 (LG1 - P)		36	XXXVI	1070	No	757	1045/357,650,38	2872
2-4* (RC - Asc)		36	XLVI	1077	No	757	1045/357,650,38	2879
2-3*		Fungtype						
4 (RC - Asc)		37	XXXVII	961	No	942	1046/357,651,38	2949
5 (RC - Asc)		38	XXXVIII	1082	268	757	1045/357,650,38	3152
13-7* (RC - Asc)		3a	XIX	No	389	No	1071/363,670,38	1460
13-8*		5c	XLIV	1082	380	No	1065/357,670,383	2527
Iceland - thermal soil								
7 (Ölkelduhals - Ast)		40	XL	1095	401	No	1071/363,670,38	2567
31 (Ölkelduhals - Ast)		1	I	1057	401	No	1071/363,670,38	2529
32 (Ölkelduhals - Ast)		1	I	1057	401	No	1071/363,670,38	2529
33 (Ölkelduhals - Ast)		40	XL	1095	401	No	1071/363,670,38	2567
34 (Reykjanes - Ast)		40	XLI	1133	401	No	1071/363,670,38	2605
35 (Reykjanes - Ast)		40	XLI	1133	401	No	1071/363,670,38	2605
36 (Reykjanes - Ast)		40	XLI	1133	401	No	1071/363,670,38	2605
37 (Reykjanes - Ast)		40	XL	1095	401	No	1071/363,670,38	2567
47 (Ölkelduhals - Ast)		40	XL	1095	401	No	1071/363,670,38	2567
49 (Reykjanes - Ast)		40	XLI	1133	401	No	1071/363,670,38	2605
Iceland – non thermal soil								
8 (Ölkelduhals - Ast)		40	XL	1095	401	No	1071/363,670,38	2567
39 (Ölkelduhals - Ast)		40	XLI	1133	401	No	1071/363,670,38	2605
41 (Ölkelduhals - Ast)		40	XLII	1134	401	No	1071/363,670,38	2606
43 (Reykjanes - Ast)		40	XLI	1133	401	No	1071/363,670,38	2605
44 (Reykjanes - Ast)		38	XLIII	1065	276	757	1046/357,651,38	3144
46 (Reykjanes - Ast)		38	XLIII	1065	276	757	1046/357,651,38	3144
48 (Ölkelduhals - Ast)		40	XLVII	1133	401	No	1071/363,670,38	2605
50 (Reykjanes - Ast)		40	XLV	1133	401	No	1071/363,670,38	2605

2. RFLP and sequence-based analyses

In this study, nine different Intra types (i.e. mtLSU RFLP patterns, Fig. 1) of *G. intraradices* could be distinguished corresponding to 15 sequence haplotypes, twelve of which were characterized for the first time. This surprisingly high number of sequence haplotypes was found because all PCR amplicons were not only analyzed by RFLP analysis, but also by sequencing. Gaining resolution beyond that of Intra types was particularly necessary for the samples from Iceland, where only three Intra types could be recorded. Intra type 40, constituting the population of *G. intraradices* in about 52% of all positive samples (Table 3) comprised five different sequence haplotypes. An RFLP pattern similar to Intra type 40 is Intra type 1, and indeed sequence analyses revealed a high similarity between the respective haplotypes although diagnostic regions allow to distinguish them (Fig. 2). To differentiate them it is sufficient to sequence the diagnostic regions in the LAGLIDADG encoding intron at position 1, using PCR primer RNL-29. Interestingly, Intra type 1 with its associated sequence haplotypes was only recorded in samples from Iceland, but not from YNP. In Iceland, Intra type 1 & 40 were dominating in the samples from thermal soils and only Intra type 38 was accompanying them in samples from non-thermal soils. Intra type 38, corresponding to sequence haplotype XLIII in Iceland, arose from the different sequence haplotype XXXVIII in non-thermal soils at YNP.

In YNP, one mtLSU sequence haplotype (haplotype XXXVII) could be found in samples both from thermal and non-thermal soils. A root sample from the non-thermal area at YNP harbored fungal strains of two RFLP types (3a; 5c), which were already known from arable fields in Tänikon and Frick (Börstler et al., 2010). The sequence of Intra type 3a was identical to the already-known haplotype XIX, however, Intra type 5c corresponded to the newly detected sequence haplotype XLIV (Table 3). All remaining four Intra types recorded throughout this study also corresponded to newly discovered sequence haplotypes.

Overall, the richness of Intra types and sequence haplotypes (number of Intra types or sequence haplotypes/number of samples; Fig. 3, Fig. 4; Fig. 5) was higher in YNP (0.77 Intra types; 0.88 haplotypes), compared to Iceland (0.16; 0.38) and also higher in non-thermal (0.42; 0.85) than thermal soils (0.38; 0.46; data pooled for Iceland and YNP).

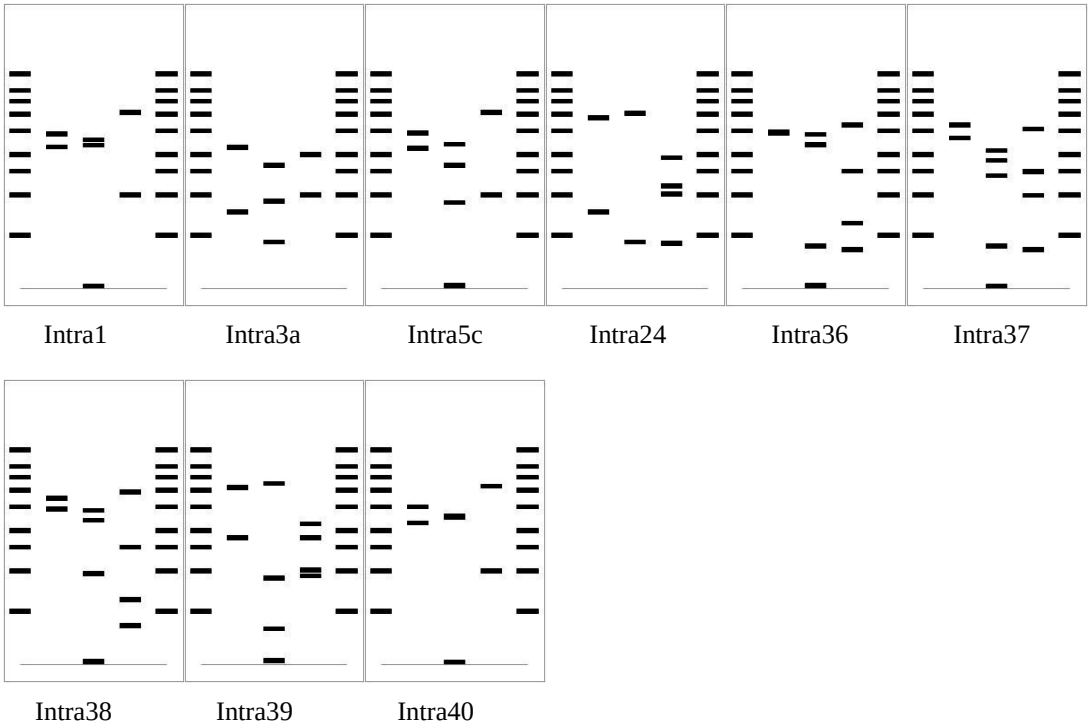


Fig. 1: Virtual restriction fragment length polymorphism (RFLP) patterns (Intra types) of the PCR amplicon of the mitochondrial rRNA large subunit gene (mtLSU) obtained with the primer pair RNL-29/RNL-30 and restriction endonucleases *Dra*III, *Bsa*JI, *Hin*dIII for phylotaxon *G. intraradices* GLOM A-1 in the present study. Fragment sizes (bp) of the DNA ladder in the flanking lanes: 250, 500, 750, 1000, 1500, 2000, 2500, 3000, 3500, 4000. The similar RFLP patterns of Intra1 and Intra40 can be distinguished by sequencing.

Sample	Region 1	Region 2	Region 3	Intra type*	Haplotype
31	GGTCTTAGA	--------AA	TGA----------	1	I
32	....	--------..	...----------	1	I
7	....	-------A..	...----------	40	XL
33	....	-------A..	...----------	40	XL
37	....	-------A..	...----------	40	XL
47	....	-------A..	...----------	40	XL
8	....	-------A..	...----------	40	XL
34	....	-------A..	...CTTTTTTAGATCGACGTAGTTCC	40	XLI
35	....	-------A..	...CTTTTTTAGATCGACGTAGTTCC	40	XLI
36	....	-------A..	...CTTTTTTAGATCGACGTAGTTCC	40	XLI
49	....	-------A..	...CTTTTTTAGATCGACGTAGTTCC	40	XLI
39	....	-------A..	...CTTTTTTAGATCGACGTAGTTCC	40	XLI
43	....	-------A..	...CTTTTTTAGATCGACGTAGTTCC	40	XLI
41	....	AAAAAGGA..	...----------	40	XLII
50	A..T...A.	-------A..	...CTTTTTTAGATCGACGTAGTTCC	40	XLV
48	....	-------A..	..CCTTTTTCGTGTAGGAATTACATC	40	XLVII

Note: * Intra-type denotes different restriction fragment polymorphisms of the PCR amplicon generated with the primers RNL-29 and RNL-30 and cleaved with the restriction endonucleases *Dra*III, *Bsa*JI, and *Hind*III (Börstler et al, 2008).

Fig. 2: Alignment of three diagnostic sequence regions of the intron at position 1, distinguishing the similar restriction fragment polymorphisms Intra1 and Intra40 (Intra types) in an mtLSU PCR amplicon. Dots indicate identical states of nucleotide positions and shaded hyphens stand for gaps. Sequence haplotype I was only found at thermal sites in this study. Haplotypes unique to non-thermal sites are highlighted by grey underlay.

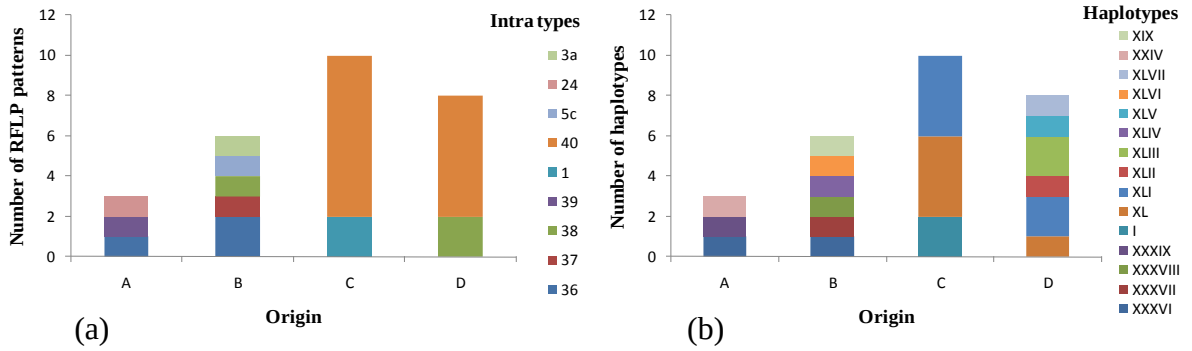


Fig. 3: Distribution of RFLP patterns and haplotypes across different sites. Number of different RFLP patterns ("Intra types") (a) and sequence haplotypes (b) of *G. intraradices* found in Yellowstone National Park (A & B) and Iceland (C & D) at thermal (A & C) and non-thermal (B & D) sites.

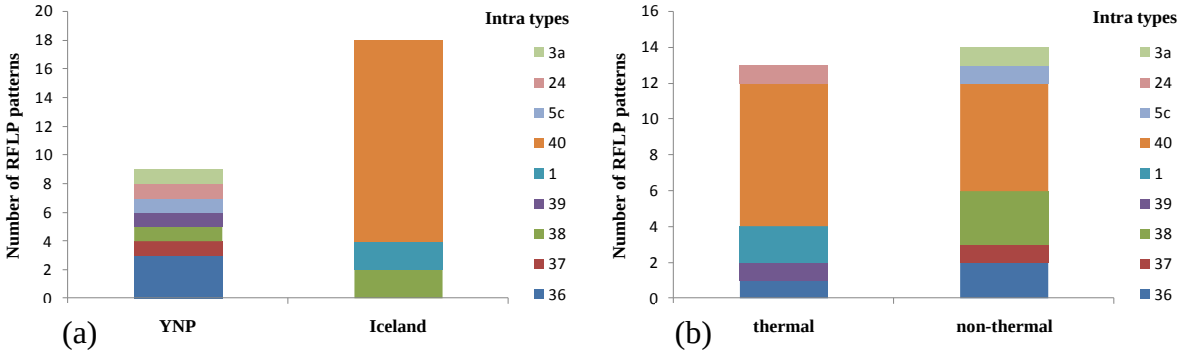


Fig. 4: Number and distribution of RFLP patterns ("Intra types") of *G. intraradices* **(a)** in Yellowstone National Park and on Iceland; **(b)** at thermal and non-thermal sites (data pooled for Iceland and YNP).

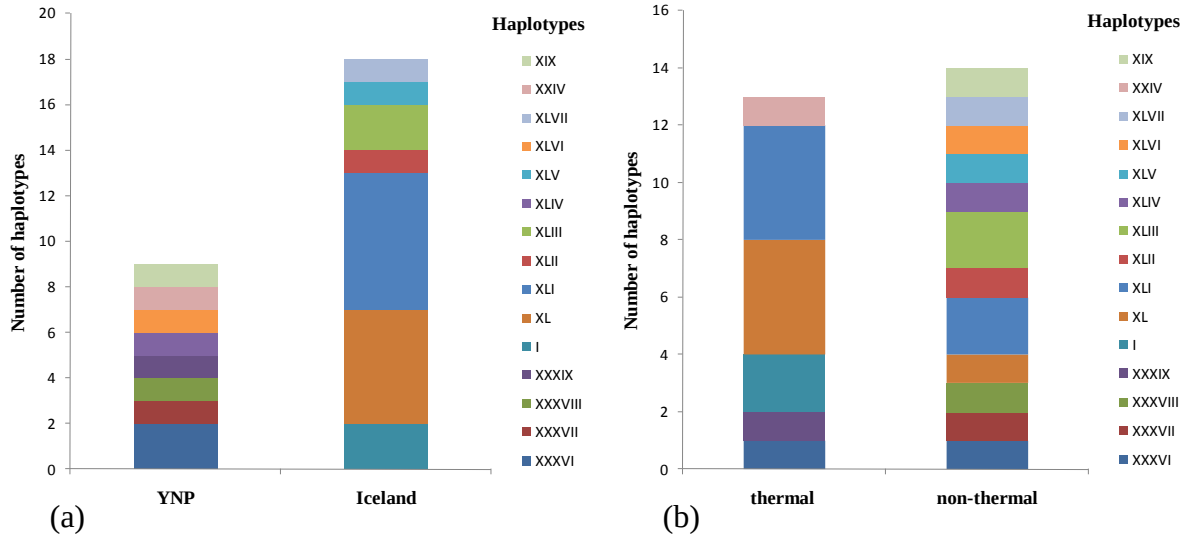


Fig. 5: Number and distribution of different mtLSU haplotypes of *G. intraradices* found **(a)** in YNP and in Iceland (data pooled from thermal and non-thermal soils), **(b)** in thermal and non-thermal sites (data pooled for Iceland and YNP).

3. Phylogenetic analyses

A phylogenetic tree was inferred from the newly generated mtLSU sequences with sequences of Raab et al. (2005) and Börstler et al. (2008, 2010) as references. The sequence haplotypes from most of the root samples from YNP fell into the 'grassland clade' consistently recovered by Börstler et al. (2010), which received maximum support in our phylogenetic analyses (Fig. 6, bootstrap value = 100%). However, none of these haplotypes from YNP falling into the "grassland clade" contained the typical intron type 2-4 (classification from Börstler et al., 2010) for all grassland haplotypes detected by Börstler et al. (2010). Two samples from YNP yielded haplotypes (XIX and XXIV), previously obtained from samples of arable field sites (Börstler et al., 2010).

In considerable contrast to the sequence haplotypes from YNP, 90% of the haplotypes from Iceland clustered with reference sequences recovered from arable fields in past surveys in Switzerland. Only two samples from non-thermal soils of Iceland harbored sequence haplotypes that together with one from non-thermal soil of YNP clustered in the 'grassland clade' and gave rise to the same Intra type 38.

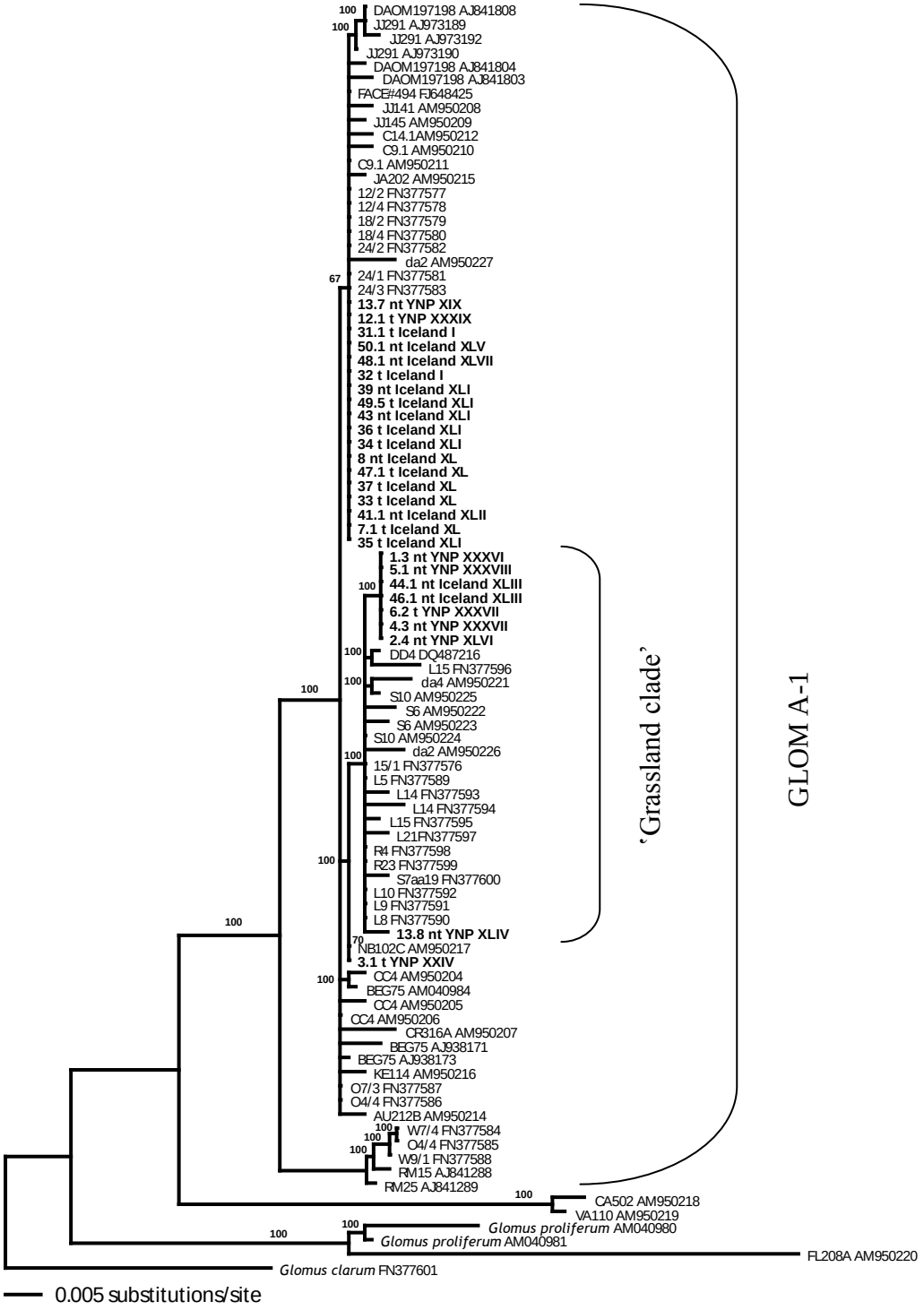


Fig. 6: Phylogenetic tree of *G. intraradices* and *G. proliferum* based on an exon mtLSU sequence of 929 bp length with *G. clarum* as outgroup. The tree was inferred by a heuristic search under the maximum likelihood criterion. The numbers at the nodes represent support values, derived from 1000 replicate maximum parsimony bootstrap inferences. Isolate code or sample code and accession number and site origin are given for previously published sequences. Sequences from this study are given in bold with indicated sample code, type of soils (t: thermal; nt: non-thermal), origin and mtLSU sequence haplotypes in roman numerals.

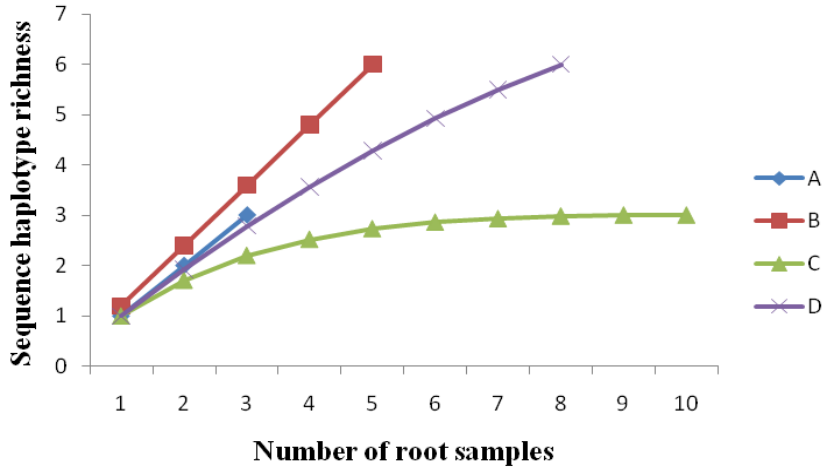


Fig. 7: Sampling effort curves for all mtLSU sequence haplotypes of *Glomus intraradices* at Yellowstone National Park (USA) in thermal (A) and non-thermal soils (B) and in Iceland at thermal (C) and non-thermal sites (D).

V. Discussion

This is the first study analyzing the diversity of mtLSU sequence haplotypes of *G. intraradices* GLOM A-1 at geothermal sites. In a previous study, on the communities of AMF in roots in sites in YNP and Iceland (Appoloni et al., 2008) *G. intraradices* was found to be the most common taxon at the sites and even occurred in some of the samples taken at the highest rhizosphere temperatures. Therefore the question arose whether intraspecific differentiation (e.g. ecotypes) might occur between thermal and non-thermal sites. At the same time, preliminary data presented by Appoloni et al. (2008) showed that species-level phylotypes were shared between YNP and Iceland, raising the possibility that specialized intraspecies genotypes are also shared by thermal areas in both locations.

Eighty percent of the mtLSU haplotypes analyzed were never found previously

Overall, the current study showed that the RFLP approach developed and optimized by Börstler et al. (2008, 2010) can be applied to basically any root sample from the field, although cloning and sequencing may be necessary to recover novel sequence haplotypes, which are otherwise not resolved by RFLP analysis. In our study, using the mtLSU PCR-RFLP-sequencing approach only 66% of the root samples from YNP tested positive for *G. intraradices*, whereas 75% of all newly prepared DNA extracts (not tested with nuclear ITS markers) from Iceland yielded an mtLSU PCR amplicon. This reduced amplification success using the DNA extracts of Appoloni et al. (2008) may be explained by poorer quality or

degradation. Many new mtLSU sequence haplotypes could be characterized (80%), some of which belong to strains of *G. intraradices* living under thermally extreme soil conditions.

MtLSU sequence haplotype richness tend to decrease in thermal sites

This study revealed that thermal and possibly physically disturbed sites by tectonic and volcanic activities tend to harbor fewer mtLSU sequence haplotypes than non-thermal, not disturbed, reference areas. This is in sharp disagreement with the finding by Börstler et al. (2010) who provided evidence that disturbance may promote genetic diversity in *G. intraradices* populations. Analyses of more samples will be necessary to confirm this preliminary evidence from our Iceland samples, where we had comparable numbers of samples from thermal and non-thermal areas. For YNP, the number of comparable samples was clearly insufficient for drawing ecological conclusions. A closer look at the geographical distribution of the mtLSU sequence haplotypes and nucleotide distribution within the sequenced mtLSU fragment suggests local and possibly also functionally relevant evolutionary radiation (see below) within *G. intraradices*.

Geographical structuring occurs between the geothermal areas of YNP and Iceland

YNP appeared to harbor a higher mtLSU sequence richness for *G. intraradices* than Iceland. However, considering that a higher number of host plant species was analyzed in YNP host preference (Sýkorová et al., 2007b) may have contributed to this impression. Apparent complete geographic differentiation between Iceland and YNP indicates dispersal limitation even in the generalist AM fungal taxon *G. intraradices*, but differences in soil chemistry and climate cannot be ruled out as contributing factors. Iceland being an island and YNP having contact to the remainder of the American continent may provide a biogeographical explanation for the finding of a reduced mtLSU richness on Iceland as compared to YNP and for the geographical structuring. More importantly, however, the finding of ever new sequence haplotypes with nearly each additionally sequenced mtLSU fragment may suggest that the local environmental conditions may sustain evolutionary differentiation and concomitant functional adaptation. The latter speculation seems justified, given that Koch et al. (2006) and Croll et al. (2008b) could show that polymorphism in mtLSU markers goes in hand with some symbiotic properties and growth parameters. Distinct distributional patterns of the mtLSU sequence haplotypes at either grassland or arable sites may reflect ecotypic differentiation as already suggested previously (Börstler et al., 2010).

However, to verify or reject these hypotheses, more habitats have to be analyzed around the world to obtain a more comprehensive picture of the genetic structure of *G. intraradices*.

An intraspecies genotype is shared among thermal soil of Iceland and arable soil

Haplotype I, detected in thermal soils of Iceland, has a widespread geographical distribution and was found particularly frequently in arable soils of Switzerland, such as those close to the villages of Tänikon, Eschikon, and Changins (Croll et al., 2008b; Börstler et al., 2010). Haplotype I is as well represented by isolate DAOM197198 from Canada, currently being used for genome sequencing (Martin et al., 2008). Interestingly, the present study identified many sequence variants from Iceland, all closely related to haplotype I.

Local genetic polymorphism is maintained in populations of *G. intraradices*

The sequences closely related to haplotype I were all highly conserved with only differences (mainly indels) in the intron at position 1, according to the nomenclature established by Börstler et al. (2010). The intron at position 1 possesses an ORF of the self-splicing mobile element LAGLIDADG 2 (Table 3; Thiéry et al., 2010). Importantly, several nucleotide polymorphisms of the mtLSU sequence haplotypes XLI, XLV and XLVIII affected the active site of the LAGLIDADG endonuclease. This is not the case for the haplotypes I, XL and XLII. Thiéry et al. (2010) gave evidence for occasional horizontal transfer of this intron at position 1 promoted by LAGLIDADG endonuclease. The present finding of nucleotide polymorphism within the active site of the LAGLIDADG endonuclease now adds to the argument that mtLSU haplotype diversity may also be of functional evolutionary importance. Indeed, homing endonucleases are mostly associated with introns and have evolved to facilitate their splicing (Burt & Koufopanou, 2004) or are involved in the long-term survival of the introns (Lucas et al., 2001). Thus, endonucleases might play an indirect role in the gene regulation by introns (Rose, 2008). Moreover, the apparent common ancestry of haplotype I and involvement of the LAGLIDADG endonuclease for the recorded swarm of mtLSU sequence haplotypes on Iceland may point at a potential mechanism promoting mtLSU sequence diversity in *G. intraradices.* At any rate, concluding from the findings of this study, processes maintaining local genetic polymorphism in populations of AMF appear to be in place. The hypothesis that a local radiation of related haplotypes occurred remains an interesting question to be further addressed in the future.

Chapter 5: General discussion

I. The development of molecular genetic markers and its contextual background

Availability of robust, sensitive, and highly resolving molecular genetic markers is crucial in many fields of molecular ecological research, such as (i) the assessment of genotypic diversity; (ii) the identification of species groups, species, individuals, and haplotypes; (iii) the estimation of fungal dispersal via studies on gene flow, and hence, in general, to address many ecological and evolutionary questions (Navajas & Navia, 2010).

In the context of AMF research, molecular markers were relevant in revealing cryptic species, undistinguishable using their spore morphologies, or to detect and identify taxa only showing hyphal growth (Gamper et al., 2010). To overcome subjectivity and the need of expert knowledge for ecological studies, DNA barcoding has been proposed (Hebert et al., 2003). DNA barcoding relies on short DNA sequences suitable for species identification that can be determined reliably for any species by using universal PCR primers (Frézal & Leblois, 2008). The DNA barcodes have to be rapidly, cost-effectively, and universally accessible. Until recently, PCR fragments of 400-800 bp in length have been proposed as ideal barcodes. A region of the cytochrome C oxidase 1 (cox1) gene, has been proposed for animals (reviewed in Begerow et al., 2010). However, for plants and fungi other possible DNA markers have been proposed due to the existence of introns in the cox1 gene. A study on the sequences of cox1 gene of 56 fungal species revealed length variation between 1,584 and 22,006 bp, mainly due to introns varying considerably in length and number (Seifert et al., 2007). The cox1 gene encompasses 11 introns in the AM fungus (AMF) *G. intraradices* FACE494 (Lee & Young, 2009). Borriello (2010) investigated the sequence diversity of cox1 in *Scutellospora*, *Gigaspora* and *Glomus* group A species, and reported low intraspecific sequence polymorphism. However, no glomeromycotan universal primers could be designed as the cox1 sequences of the various AMF lineages appear to be too divergent (Borriello, 2010). Altogether this shows that cox1 is rather unsuitable as a marker gene for molecular genetic barcoding in AMF. Instead, a fragment of 1500 bp spanning the nuclear SSU, ITS and LSU was proposed by Stockinger et al. (2010) as a barcoding region for AMF, as shorter nuclear rDNA sequences do not resolve closely related species.

This PhD project aimed at identifying and characterizing DNA regions, suitable as molecular genetic markers at the sub-species level of detection and identification. These markers were intended for studying natural populations of AMF in different ecosystems. A previous MSc thesis by Anna Ochsner (2002) and the PhD thesis by Philipp Raab (2007) indicated the great potential that markers based on the mtLSU gene of AMF may have. Raab et al. (2005) published the first mtLSU sequences of two *G. intraradices* isolates, one *G. proliferum* isolate and an environmental root sample. He showed that within the same fungal, there is no sequence polymorphism. This homoplasmy of the mitochondrial genome greatly facilitates molecular genetic detection and identification and thus circumvents the problems with nuclear encoded DNA markers caused by sequence heterogeneity. Börstler et al. (2008) confirmed the homogeneity of the mtLSU within AMF isolates and developed a PCR-RFLP approach to differentiate *G. intraradices* haplotypes.

In chapter 2 of this thesis project, I explored the resolution of the mtLSU in other AMF species, using the sequence data of *G. intraradices* (Raab et al., 2005; Börstler et al., 2008). In chapter 3, additional DNA regions of mitochondrial intergenic spacers were examined for their possible suitability as molecular genetic markers for population studies. Chapter 4 deals with the application of the PCR-RFLP approach (Börstler et al., 2008) on environmental root samples from geothermal and non-geothermal sites, as it is so far the only molecular genetic approach for population-level studies in roots.

II. mtLSU as a molecular marker in AMF

Ecological investigations into the effect of environmental factors on populations and communities of AMF have been hampered by the lack of universal criteria to define and recognize phylotypes (Öpik et al., 2010) and by pronounced sequence diversity in the nuclear rDNA (Stockinger et al., 2010). Recent methodological development and first field investigations indicated that these limitations for molecular genetic studies can be avoided by the use of markers targeting the mtLSU gene. The mtLSU gene turned out as a highly suitable source for developing molecular genetic markers because of the existence of discriminative regions and most importantly the genetic uniformity of the mitochondrial genome in AMF (chapter 2; Börstler et al., 2008). Börstler et al. (2008) could resolve 12 haplotypes out of 16 isolates using the mtLSU. In comparison, Mathimaran et al. (2008a) identified 18 polymorphic simple sequence repeat (SSR) markers that could resolve 7 genotypes out of 8

isolates and Croll et al. (2008b) could resolve 18 genotypes out of 48 isolates using SSR, nuclear gene intron and mtLSU gene intron markers. Mathimaran et al. (2008a) and Croll et al. (2008b) differentiated *G. intraradices* isolate JJ291 from isolate DAOM197198, indicating that the markers they used give a better resolution that the mtLSU. However, the resolving power of the different molecular markers is difficult to compare as separate case studies shared only a few isolates. The mtLSU RFLP types and associated sequence haplotypes defined by Börstler et al. (2008, 2010) have the considerable advantage over microsatellites (Croll et al., 2008a) that they originate from known regions and contain sufficient phylogenetic information for discrimination from any potential co-amplified non-target organism. The possibility to perform such methodological validation checks is indispensable, if markers are to be applied on environmental samples. Last but not least, it is the grouping by mtLSU phylotypes that gave best correlations with functional attributes of AMF (Koch et al., 2006; Croll et al., 2008b).

Yet, when this PhD thesis project was started, the sequence data for the mtLSU was restricted to two species, *G. intraradices* and *G. proliferum*. Sequencing of PCR amplicons of five additional species of *Glomus* group A (*G. mosseae*, *G. geosporum*, *G. caledonium*, *G. clarum*, *G. coronatum*), a species from *Glomus* group B (*Glomus etunicatum*-like), and two species from Gigasporaceae (*Scutellospora verrucosa* and *Scutellospora castanea*) broadened taxonomic coverage. As described in chapter 2, the differences in the mtLSU sequences were found to be considerably less among and within these other species, compared to *G. intraradices*: (i) two isolates of *G. clarum* and four isolates of *G. mosseae* did not show any mtLSU sequence differences, (ii) the congeneric sets of species, *G. mosseae*, *G. coronatum*, *G. geosporum* and *G. caledonium*, and *S. verrucosa* and *S. castanea*, both displayed little polymorphism within the nucleotide sequences.

Use of rRNA genes as targets for molecular genetic markers is adequate to study both active and inactive fungi (Anderson & Parkin, 2007). Detection of fungi in the symbiotic state was possible through the PCR-RFLP approach developed by Börstler et al. (2008). The PCR-RFLP assays were first applied to environmental root samples by Börstler et al. (2010; appendix 2) and were further used for the field investigation described in chapter 4. The fact that the number of mitochondria increases with physiological activity (Besserer et al., 2006) could be used for activity measurements via quantitative PCR. The possibility to work on DNA could be a particular advantage for molecular ecological studies, because this would

avoid the complications associated with the inherent instability of RNA, upon which activity measurements are usually based on.

III. Sequence evolution in the mtLSU gene

This PhD thesis provides first insights into the evolutionary dynamics of introns and HEs within the mtLSU in the phylum Glomeromycota in chapter 2. AMF species and many isolates of *G. intraradices* differed strongly with regard to their introns. Contrary to intron 1513 that has apparently been inherited strictly vertically from the ancestors of *Glomus* groups A and B, introns position 1149, 1187, 1450 and 2184 of the mtLSU (positions in the *G. intraradices* exon FACE494 (Lee & Young, 2009)) encompassed HE ORFs (chapter 2, appendix 1). Generally, HE ORFs are assumed to follow a ‘‘homing cycle”, that is fixation within a gene, degeneration, intron loss and possible ‘‘re-colonization” to ensure their propagation in introns (Goddard & Burt, 1999; Fig. 1; chapter 1, section VII. 2.). In this thesis, we demonstrated for the first time that the HEG go through major stages of the ‘‘homing cycle” in the Glomeromycota.

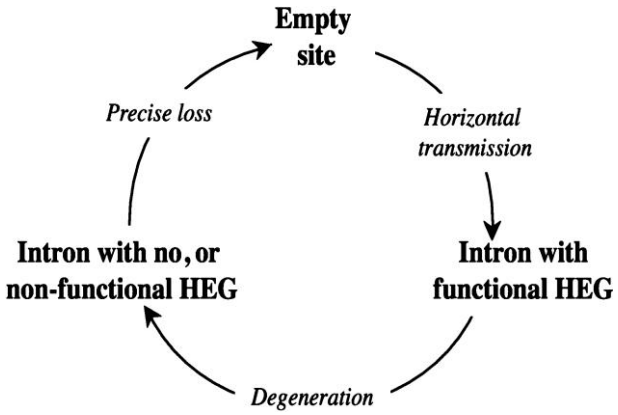


Fig. 1: Cycle of intron gain and loss: an empty site (a gene without intron) is invaded by a homing endonuclease gene (HEG) through horizontal transmission, the HEG degenerates, eventually the intron is lost, and the gene is reinvaded. From Sandegren & Sjoberg (2004).

Intron 1149 of one isolate of *G. intraradices* showed a striking similarity to isolates of *G. clarum*, and provided evidence for horizontal transfer of the intron, either from *G. intraradices* to *G. clarum* or vice versa. Phylogenetic analyses suggested a HEG spread or a co-evolution within the true fungi by horizontal transfer. A possible way of transfer of the HEGs might be possible through anastomoses. In contrast to the study of Giovannetti et al. (2004) and Voets et al. (2006) who observed anastomoses only within the same isolate for *G.*

mosseae and *G. intraradices*, Croll et al. (2009) demonstrated that genetic exchange through anastomoses was possible among different isolates of *G. intraradices* and not restricted to genetically identical organisms.

IV. Exploration of additional possible mitochondrial molecular genetic markers

Our data suggest mitochondrial homoplasmy, i.e. all copies of mtDNA within an AMF individual have the same sequence, in agreement with the studies of Raab et al. (2005) and Börstler et al. (2008). The haploid nature of mitochondria makes mitochondrial genome sequences highly-sensitive molecular markers with no cloning steps required. Within the mitochondrion, different types of molecular markers could be designed to achieve diverse tasks (e.g. discriminating strains, species or inferring phylogeny). Non-coding DNA displays polymorphism that is commonly used for population studies. This in mind, four mitochondrial intergenic spacers were assessed for their polymorphism among six isolates of *G. intraradices* differentiated by their mtLSU haplotype, a field sample of known mtLSU haplotype and database sequences from two additional isolates in chapter 3. Combination of the four investigated intergenic spacers could allow distinction among isolates not resolved using the mtLSU haplotypes. The rns/nad5 spacer, a DNA region comprising HE ORFs, displayed the highest resolution. It did not come as a surprise as we showed polymorphism in HE ORFs due to the "homing cycle" (chapter 2; chapter 5, section I.3.). That opens up new perspectives for genotyping approaches at the population-level and future research should focus on such rapidly-evolving genomic regions.

V. mtLSU sequence diversity in populations of *G. intraradices* in extreme environments

At geothermal sites, plants were found with different genotypes possibly corresponding to ecotypes (e.g. *Agrostis scabra* occurs in two ecotypes depending on its growth on thermal or non-thermal sites) and they have developed diverse strategies in extreme environments to cope with elevated temperature: (i) growing on moss mats for isolation and to retain moisture (Tercek & Whitbeck, 2004); (ii) forming more, shorter and highly branch roots; (iii) producing heat shock proteins (Stout et al., 1997) as a short-term reaction; (iv) using water and steam resources from their environment; (v) forming symbiosis. In this PhD thesis, we were interested in knowing the degree of intraspecific diversity of the most probably widespread fungus *G. intraradices* hypothesizing that they may confer adaptive advantages to their host plants under extreme environmental conditions. To achieve our goal,

we used the mtLSU PCR-RFLP approach developed by Börstler et al. (2008), the only molecular marker applied under field conditions for population studies of *G. intraradices*.

Appoloni et al. (2008) showed that geothermal soil at Yellowstone National Park (YNP) harbored distinct AMF community composition with both unique phylotypes and generalist fungi, with *G. intraradices* GLOMA1 the most frequent phylotype detected. In this PhD thesis, strong biogeographical structuring in the occurrence of *G. intraradices* strains was reported in roots at geothermal and non-geothermal sites of YNP and Iceland. The mechanisms to maintain this intraspecific diversity are still unknown. We suspect contribution of HEs in this process as we found nucleotide polymorphism within its active site. We hypothesized HEs to play an indirect role in gene regulation through introns by being involved in intron splicing (Burt & Koufopanou, 2004) or intron long-term survival (Lucas et al., 2001). Two hypotheses have been put forward that explain the occurrence of intraspecific diversity (Rosendahl & Matzen, 2008): (i) diverse microhabitats with niches can originate from an apparent homogenous soil, (ii) isolation by distance. The entire life-cycle of AMF is completed below ground, limiting then the possibilities for dispersal. However, some haplotypes of AMF have a global distribution and might be linken to ecological generalist with high dispersal abilities, e.g. haplotype I was recorded in arable fields in Switzerland (Börstler et al., 2010) and at thermal sites in Iceland. Its ruderality and adaptation to physical disturbance may be attributed to its ability to exploit a range of diverse resources, notably in the case of habitat quality changes or habitat loss caused by disturbance (e.g. agricultural or volcanic activity), contrary to other haplotypes that might not be able to adapt to new conditions.

Almost every sample analyzed at geothermal sites generated new mtLSU haplotypes. This suggests considerable genotypic richness of *G. intraradices* under various environmental conditions and a maximum of sites should be analyzed to cover *G. intraradices* diversity. Bunn et al. (2009) demonstrated that AMF from thermal soils were not specifically adapted to high temperatures as the source of inoculum from thermal or non-thermal soils had no impact in terms of colonization effect, extraradical hyphae length, host plant biomass, flowering for host species in artificially elevated-temperature soils. Instead, Lekberg et al. (2007) underlined the importance of soil characteristics (composition, pH) and spatial structure in shaping AMF communities, suggesting that AMF species or AMF populations account more for these environmental variable that the influence of temperature.

VI. Conclusions and perspectives

This PhD thesis project gave first insight into the potential that mitochondrial genes and intergenic spacers have in marker development for molecular ecological studies of AMF ecology and evolution at lower taxonomic levels. Four main conclusions could be drawn (Fig. 2):

i) The mtLSU gene harbours suitable intron regions varying in length and nucleotide composition for studies on sequence diversity at the species level.

ii) Whereas some introns of the mtLSU gene appear to have been inherited strictly vertically from the common ancestor of the *Glomus* groups A and B, others containing ORFs of HE, seem to be transferred among taxa by occasional horizontal gene transfer. Evolutionary sequence dynamics after such events of horizontal gene transfer involve intron retention, degeneration, as well as loss, as observed in the form of ORFs of HEs at different stages of integrity.

iii) The combined nucleotide polymorphism in four mitochondrial intergenic spacers, namely rns/nad5, atp6/nad2, nad3/nad6, and nad6/cox3 turned out to provide better resolution than the sequence polymorphism in the single mtLSU gene marker. The rns/nad5 spacer, a spacer encoding an ORF of a HE, displayed the highest resolution for different isolates of AMF, possibly due to the associated evolutionary sequence variation.

iv) The molecular ecological survey at geothermal sites in Europe and North America recovered many new mtLSU sequence haplotypes of *G. intraradices* along with some that have been recorded previously in roots from arable fields (Börstler et al., 2010). In fact, the analysis revealed evidence for considerable biogeographical isolation in the occurrence of many mtLSU sequence haplotypes of *G. intraradices*.

Possible perspectives arising from the findings of this PhD project could be:

i) Investigations into the relationship between the diversity of AMF strains (mtLSU sequence haplotypes) and mycorrhizal symbiotic functioning, as for instance measured in terms of plant growth, mineral nutrition, and biotic and abiotic stress alleviation. To this end, inoculation trials would be needed with each AMF isolates and species for which the mtLSU

sequences have been characterized *G. geosporum*, *G. coronatum*, *G. caledonium*, *G. clarum*, *G. mosseae*, *G. intraradices*, *S. verrucosa*, *S. castanea* and *G. etunicatum*-like (chapter 2).

ii) Further investigations into DNA sequence evolution within and among AMF species to follow up the finding of possible horizontal gene transfer between *G. intraradices* and *G. clarum* (chapter 2). Further research is needed in other species and experimental tests are necessary to address whether genetic exchange is positively correlated with anastomosis formation. Particularly promising seem experimental pairings of two AMF strains, one containing a gene of a HE and one lacking it.

iii) By characterizing further mitochondrial DNA sequences as the work described in chapter 3, it should now become possible to study mitochondrial inheritance and the mechanism leading to the maintenance of a genetically uniform mitochondrion.

iv) Studies on the physiological activity of AMF should now become possible, given that we have now AMF-specific PCR primers (chapters 2 and 3) and given that the number of mitochondria has been shown to increase with fungal physiological activity (Besserer et al., 2006). Assays for quantitative PCR could be developed targeted to some of the newly characterized mitochondrial DNA regions.

v) Experiments to test the effect of various environmental parameters on the composition and structuring of populations of *G. intraradices*, would be promising, given evidence of strong biogeographical structuring (chapter 4). Variation of the experimental parameters could promote selective growth of different ecotypes of *G. intraradices*, the mtLSU sequence haplotypes of which could then be detected and identified.

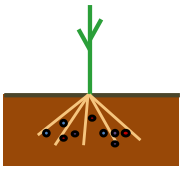
vi) Moreover, it would be interesting to explore the possibility of horizontal gene transfer between AMF and their host plants. The intimate association between plant and fungus in the AM symbiosis may make this possible, given that nuclear horizontal gene transfer has been observed between the parasitic plant *Striga hermonthica* and its monocot host plants *Sorghum bicolor* and *Oryza sativa* (Yoshida et al., 2010). A similar mechanisms of horizontal gene transfer as that proposed for these pairs of plants could be imagined for AMF-plant pairings as well. The suggested mechanism of horizontal gene transfer is that mRNA is translocated over the haustorial interface in the plant-plant pairings, which may happen via the arbuscular or

intraradical mycelial interface in the AMF-plant symbiosis. In fact, this has already been proposed based on high sequence similarity between introns of the AMF fungus *Glomus diaphanum* and plant pathogenic fungus *Rhizopus oryzae* and its host plants (Vaughan *et al.*, 1995; Lang & Hijri, 2009).

Indeed, this PhD thesis project showed the promising nature of the genetic characteristics of the AMF mitochondrion for development of molecular genetic markers and evolutionary sequence analyses. It is anticipated that much will be learned soon about the biology, ecology, and evolution of AMF via use of properties of mitochondrial nucleotide sequences.

THIS THESIS:

General introduction (Chapter 1)



Investigation of the mtLSU in several AMF species (Chapter 2)

Insights about evolutionary dynamics of introns and HE ORFs (Chapter 2)

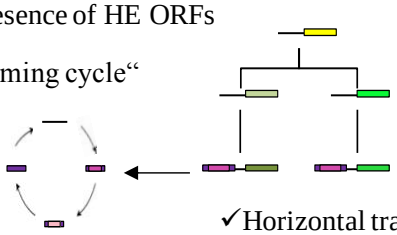
Main findings:

mtLSU + intergenic spacers

Mitochondrial genome

✓Presence of HE ORFs

„Homing cycle“



✓Horizontal transfer via anastomosis?

✓Nucleotide polymorphism

Inter-isolate polymorphism of intergenic spacers in *Glomus intraradices* (Chapter 3)

Molecular genetic markers for population & community studies on AMF (Chapters 2 & 3)

Main findings:

✓Presence of HE ORFs
✓Nucleotide polymorphism

mtLSU marker

Distribution of mtLSU sequence haplotypes of *G. intraradices* at geothermal & non-geothermal sites (Chapter 4)

Main findings:

✓Newly-discovered + previously-known haplotypes

✓Biogeographical structuring

Perspectives:

Relationship between mtLSU sequence haplotype diversity & ecological functions of AMF

Mechanisms of mitochondrial inheritance & maintenance of mitochondrial homoplasmy

Fig. 2: Recapitulative scheme of the PhD thesis studies. In red the objectives of the PhD thesis are described; in black and ticked off are the main findings; and in blue the perspectives are given.

Acknowlegments

I would like to express my sincere gratitude:

- To my supervisor Prof. Dirk Redecker for having given me the opportunity to conduct my PhD in the "Molecular Ecology of Arbuscular Mycorrhizal Fungi" group of Basel, for proposing this motivating project, for his guidance, support, constructive feedback and inspiring suggestions.

- To Prof. Andres Wiemken and Prof. Thomas Boller for welcoming me in the botanical institute and for their constant support throughout my PhD thesis.

- To the former members of my group: Boris Börstler for his daily partnership in the lab, encouragement, for sharing his knowledge and leading stimulating discussions; for all his efforts in the characterization of mtLSU haplotypes of *Glomus clarum*, and for providing me DNA of *G. intraradices*. Zuzana Sýkorová for her warm welcome in the group, her kindness and correcting my introduction; and both Boris and Zuzana for their nice introduction to lab work; Philipp Raab and Susann Appoloni for their cordial support and for providing me the starting point of my projects. Thanks to all of them for the good atmosphere in the lab and the fun we had together.

- To Hannes Gamper for spending time to read and correct the chapters 2, 3, 4 and 5 of my thesis, and for all his nice suggestions. To Kurt Ineichen for providing and cultivating the AMF and for all his help. To Natarajan Mathimaran for providing me a first DNA extract of *S. verrucosa*. To all the other former and present members of the "mycorrhiza" group in particular Mohamed Al-Yahya'ei, Yi Song, Sally Koegel, Sarah Symanczik, Pierre-Emmanuel Courty, Emilia Pelc, Florian Walder, Sietse van der Linde, Stefanie Burger, Atti Tchabi and also to the other colleagues of the institute including Deborah Strebel, Tobias Mentzel and Birgit Schulze. I want to thank them for interesting discussions, their enjoyable company and great time we spent together.

- To Václav Mandák, Erika Roth, Jürg Oetiker and Giacomo Busco for their help in making everything working smoothly.

- To co-workers who helped to sample roots from Yellowstone National Park and Iceland: Catherine Zabinski, Rebecca Bunn, Lorna McIntyre, Ylva Lekberg, Mike Tercek (Montana State University, Bozeman, USA); Hafdís Hanna Ægisdóttir (Botanical Institute, University of Basel, Switzerland) and Ulfur Oskarsson (Soil Conservation Service, Gunnarsholt, Iceland).

- To my entire family, especially to my parents and siblings: Shirley and Jean-François, Claire, Jeanne, Anne, Marie, Georges, Thérèse and Joseph. Thank you for their unconditional love, patience and permanent support.

- To the Swiss National Science Foundation for financial support by the grant 3100A0-109466.

Appendix 1: Genetic diversity of the arbuscular mycorrhizal fungus *Glomus intraradices* as determined by mitochondrial large subunit rRNA gene sequences is considerably higher than previously expected

Published in New Phytologist [2008] **180**: 452–465

Boris Börstler, Philipp Raab, Odile Thiéry, Joseph B. Morton and Dirk Redecker

I. Summary

- *Glomus intraradices* is a widespread arbuscular mycorrhizal fungus (AMF), which has been found in an extremely broad range of habitats, indicating a high tolerance for environmental factors and a generalist life history strategy. Despite this ecological versatility, not much is known about the genetic diversity of this fungal species across different habitats or over large geographic scales.
- A nested polymerase chain reaction (PCR) approach for the mitochondrial rRNA large subunit gene (mtLSU), distinguished different haplotypes among cultivated isolates of *G. intraradices* and within mycorrhizal root samples from the field.
- From analysis of 16 isolates of this species originating from five continents, 12 mitochondrial haplotypes were distinguished. Five additional mtLSU haplotypes were detected in field-collected mycorrhizal roots. Some introns in the mtLSU region appear to be stable over years of cultivation and are ancestral to the *G. intraradices* clade.
- Genetic diversity within *G. intraradices* is substantially higher than previously thought, although some mtLSU haplotypes are widespread. A restriction fragment length polymorphism approach also was developed to distinguish mtLSU haplotypes without sequencing. Using this molecular tool, intraspecific genetic variation of an AMF species can be studied directly in field plants.

Key words: arbuscular mycorrhiza, Glomeromycota, *Glomus intraradices*, intraspecific diversity, mitochondrial haplotypes, molecular markers

II. Introduction

Arbuscular mycorrhizal fungi (AMF) are associated with the broad majority of plant species and play an important role in mineral nutrient uptake. In exchange for photosynthates provided by the plant symbionts, the fungal partners improve the plants' access to phosphate, nitrogen and other mineral nutrients. The diversity of AMF correlates with diversity of plant communities suggesting that AMF influence competitive interactions among plants (Streitwolf-Engel et al., 1997; van der Heijden et al., 1998).

Molecular methods have been developed that allow identification of AMF within roots without the necessity of spore formation. In all studies to date addressing genetic diversity of AMF in roots in the field, only regions of nuclear-encoded ribosomal RNA genes have been used. Specific polymerase chain reaction (PCR) primers amplify diagnostic regions of these genes from colonized roots (Redecker, 2006). The resulting PCR products are characterized by various methods, including restriction fragment length polymorphism (RFLP) and DNA sequencing to identify the fungi. The applications of molecular identification methods in field settings have yielded novel insights into the ecology of these fungi (Öpik et al., 2006).

When using nuclear rRNA genes for phylogeny and identification of glomeromycotan fungi, the high variation among gene copies, present even within single spores of these organisms (Sanders et al., 1995; Lloyd MacGilp et al., 1996; Lanfranco et al., 1999) impairs not only the identification of closely related morphospecies, but also differentiation of isolates within a morphospecies. Variation is more acute in the internal transcribed spacers (ITS) than in the more conserved regions of rRNA genes. For example, ITS sequences within a single spore isolate of *Glomus intraradices* were as divergent as sequences from other isolates (Jansa et al., 2002b).

Intraspore rRNA gene variation could occur among rRNA gene copies in the genome of a single nucleus, as reported from other organisms (Buckler et al., 1997) or among nuclei inhabiting the same cell. The genetics of multiple nuclei in the glomeromycotan mycelia is conflicting, with some evidence suggesting nuclear populations are heterokaryotic (Kuhn et al., 2001; Hijri & Sanders, 2005) and other data indicating they are homokaryotic (Pawlowska & Taylor, 2004). A heterokaryotic genetic system implies absence of a fixed nuclear genotype for a fungal isolate, with populations of nuclei changing within a species. Rosendahl (2008) summarized recent progress in the budding field of AMF population biology.

Koch et al. (2004) characterized isolates of *G. intraradices* cultivated in root organ cultures using amplified fragment length polymorphism (AFLP), showing a high degree of

genetic and phenotypic diversity among those isolates. With the exception of the Canadian isolate DAOM197198, all isolates originated from one field site in Tänikon, Switzerland (Jansa et al., 2002b).

Croll et al. (2008b) used a larger set of root organ cultures of *G. intraradices* isolates from the same field site to elucidate local genetic diversity. They used 10 simple sequence repeat (SSR) loci as molecular markers as well as introns of the mitochondrial large subunit rRNA gene (mtLSU; Raab *et al.* 2005) and introns of a nuclear gene. Genetic diversity among fungal isolates was high, but isolates from other locations in Switzerland and Canada were not substantially different. These results complemented those of Koch et al. (2004) and indicated that much of the global genetic diversity of *G. intraradices* could be represented just within this field site. Multilocus genotypes also have been identified from individual fieldcollected spores of *G. mosseae* using markers from single copy nuclear genes *GmFOX2*, *GmTOR2*, and *GmGIN1* (Stukenbrock & Rosendahl, 2005b).

Mathimaran et al. (2008a) used a set of 18 SSR markers to analyse genetic variation among eight isolates of *G. intraradices.* Only two isolates from this set appeared to be identical clones. Neither the SSR or AFLP markers listed earlier, nor the 'Single Nucleotide Polymorphisms' of Stukenbrock & Rosendahl (2005b) have been applied to mycorrhizal roots from the field so far.

Mitochondrial DNA has a long history as a molecular marker (especially for the Metazoa), which preceedes the era in which PCR facilitated the access to its sequences from a broad range of organisms (Bruns et al., 1989). A region of the mtLSU was used so successfully for routine molecular identification of ectomycorrhizal fungal species from colonized roots that a large dataset is available for comparative study (Bruns et al., 1998).

Use of mitochondrial genes avoids possible complications from heterogeneous sequences encountered with nuclear genes. Raab et al. (2005) provided the first sequences from the mitochondrial genome of the Glomeromycota and documented the absence of any substantial variation in an mtLSU region within isolates of *G. intraradices* and *G. proliferum*. However, sequences were polymorphic among isolates of this species. Most notable was the presence/absence of introns and sequence variation within introns. These results suggest that mtLSU sequences provide useful information that distinguishes closely related *Glomus* species as well as intraspecific variation. A practical aspect of this approach is using mtLSU data to determine haplotypes of fungal symbionts directly amplified from mycorrhizal roots, an essential criterion for population studies of nonculturable organisms. Specific primers have

been designed which directly amplify mtLSU sequences from mycorrhizal roots (Raab et al., 2005). Molecular analyses of field-collected mycorrhizal roots reveal high diversity and putative taxa that do not sporulate (Helgason et al., 2002). However, sequence types corresponding to a few well-known morphospecies were also detected in a broad range of habitats. *Glomus intraradices* has been the most common species detected in a range of studies and it is one of the most extensively studied species in Glomeromycota. This species was found in mycorrhizal roots in habitats as different as high-input and low-input agricultural field sites (Hijri et al., 2006) and species rich grasslands (Sýkorová et al., 2007b) in Switzerland, phosphate-polluted sites (Renker et al., 2005) and mountain meadows (Börstler et al., 2006) in Germany, and geothermal soils in Yellowstone National Park, USA (Appoloni et al., 2008). All of this ITS sequence variation clustered phylogenetically within a clade of sequences originating from a single spore of *G. intraradices* (Jansa et al., 2002b). Using more conserved rRNA gene regions *G. intraradices* was detected in grasslands of Estonia (Öpik et al., 2003) and even in tropical trees in Panama (Husband et al., 2002). This fungal species has been classified as a generalist because it is abundant across disturbed as well as more mature habitats (Sýkorová et al., 2007a). It is widespread geographically and tolerates a wide range of habitats (Öpik et al., 2006). It is also compatible with all culturing systems currently in use, from glasshouse pots to root-organ cultures (Jansa et al., 2002a), and thus is one of the most common fungal components in commercial inocula (Corkidi et al., 2004). Not surprisingly, therefore, it was chosen as the model AMF species for genome sequencing (Martin et al., 2004). Given the importance and ubiquity of this species, a detailed understanding of population structure is essential.

Defining boundary conditions for *G. intraradices* has been problematic because variation in morphological features intergrades within and between isolates. Spore wall organization and structure is conserved and diagnostic, but number and color of layers are variable so that spore populations can vary considerably in size and color. Also, isolates can vary greatly in frequency and degree of aggregation in roots and/or soil.

This study addressed the following questions: Are mtLSU sequences polymorphic among isolates from different geographic locations? Do intron sequences provide stable markers that distinguish fungal haplotypes? An applied outcome of this work was an easy-to-use genotyping system based on mtLSU markers to study intraspecific genetic variation in the field.

III. Materials and Methods

1. Root organ cultures of *G. intraradices*

Isolate CC-4, originating from a fallow field in Clarence Creek, Ontario, Canada, was purchased from the Glomeromycota *in vitro* Collection (GINCO)/Belgium (ID codes MUCL43204, DAOM229456; for details see http://emma.agro.ucl.ac.be/ginco-bel/index.php). Isolate DAOM197198 originated from Pont Rouge, Québec, Canada, tree plantation/*Fraxinus americana* and was obtained independently from G. Bécard (University of Toulouse, France) in 1995 and from N. Requena (University of Karlsruhe, Germany) in 2005. This isolate also is known under the ID codes MUCL43194 and DAOM181602. Isolates JJ141, JJ145, and JJ183 originated from Hausweid, Tänikon, Switzerland (long-term field tillage experiment including crop rotation), and were obtained from J. Jansa (Jansa et al., 2002b). All isolates were propagated in root organ cultures (ROCs; see Fig. 1) on transformed carrot roots as previously described by Bécard & Fortin (1988). For DNA extraction, spores (see Fig. 2) were dissolved in 10 mM sodium acetate–citrate buffer (pH 6.0) and washed in sterile water according to Doner & Bécard (1991). Croll et al. (2008b) and Koch et al. (2004) used the isolate codes B7, C5, C2 and C3 for JJ291 (Raab et al., 2005), JJ141, JJ145 and JJ183, respectively.

2. Inocula and pot cultures of *G. intraradices*

Isolates of *G. intraradices* from the International Culture Collection of (Vesicular) Arbuscular Mycorrhizal Fungi (INVAM) were obtained as pot culture substrate-inoculum (see http:// invam.caf.wvu.edu/index.html): AU212B (Australia), CA502 (California/USA), CR316A (Costa Rica), FL208A (Florida/USA), JA202 (Kitami Agricultural Station/Hokkaido/Japan from *Phaseolus vulgaris* crop), KE114 (Kenya), NB102C (Namibia; from a native bush in the Namib desert), SW205 (Switzerland; same strain as JJ141 contributed to INVAM, for details see ROCs in section 2.3.1) and VA110 (Virginia/USA; soil from suburban home garden near Washington, DC). Isolate DD-4 (accession number DQ487216, nuclear-encoded small subunit rRNA gene) originated from a Dutch dry dune grassland (Provinciale Waterleidingduinen/Netherlands; 52°36′N, 4°38′E) and was obtained from M. G. A. van der Heijden (Vrije Universiteit Amsterdam) as pot culture substrate-inoculum. All pot culture substrate-inocula were stored at 4°C until DNA extraction. For

isolate DD-4, a new pot culture (Fig. 3) was set up and cultured under the greenhouse conditions described in Tchabi *et al.* (2008). Substrate consisted of sterilized Terragreen (American aluminium oxide, Oil Dry US special, Type III R, < 0.125 mm; Lobbe Umwelttechnik, Iserlohn, Germany) and Loess from a local site mixed 9:1 (w:w). Approximately 1 g of contributed inoculum was layered under seeds of *Plantago lanceolata* and *Hieracium pilosella*. After 8 months of cultivation, a 10–15 ml sample from the original pot culture substrate-inocula and from the new culture of DD-4 was wet-sieved using a top sieve with 1 mm openings and a bottom sieve with 38 µm openings. The content of the bottom fraction was collected in 20 ml water, applied to a 70% (w:v) sucrose solution and centrifuged at 820 *g* for 2 min (Esch et al., 1994; see Fig. 4). All organic matter suspended in the supernatant was decanted, repeatedly rinsed in a 38 µm sieve, transferred to 1.5 ml tubes, and stored on ice until DNA extraction. Single spores were collected separately, washed thoroughly in distilled water and placed in 0.2 ml tubes.

3. Field-collected root samples

The DNA extracts of mycorrhizal roots collected from the 'Ramosch'-meadow in the Engadin region of Switzerland were provided by Z. Sýkorová (Sýkorová et al., 2007b). Samples S6 (*Trifolium* sp.) and S10 (*Trifolium* sp.) correspond to the samples 11-2v and 11-3a, respectively, of Sýkorová et al. (2007b) in the Supporting Information, Table S1. DNA extracts from samples of two *Plantago lanceolata* mycorrhizal root systems originating from the Gyöngyösoroszi mine spoil in the Matra mountains of Hungary (47°50′44″N, 19°53′05″E) were provided by I. Parádi (Eötvös Loránd University, Budapest, Hungary). These two samples were designated da2 and da4.

4. DNA extraction

The DNA of 1–15 spores was extracted according to Redecker et al. (1997): spores were crushed in 2 µl 0.25 N NaOH, heated to 95°C, and incubated for 2 min. One microliter of 0.5 M TrisHCl (pH 8.0) and 2 µl 0.25 N HCl were added to this crude extract, which was heated again for 2 min and then used directly as PCR template. DNA was extracted from: spore populations (> 15); 50–80 mg (wet weight) of the organic fraction extracted from

inocula/pot culture substrates and 50–80 mg (wet weight) of plant roots using the DNeasy plant mini kit (Qiagen, Hilden, Germany). Liquid nitrogen was used to grind frozen root samples. Depending on success of amplification, extracts were further diluted 1:10 and 1:100 in TE (Tris-ethylenediaminetetraacetic acid) buffer and used again as PCR template (Table 1).

5. PCR amplification of mtLSU

Isolates CC-4, DAOM197198 and JJ183 were amplified by nested PCR as described in Raab *et al*. (2005) with slight changes. For the first reaction the primer pair consisted of RNL-3 and RNL-9. Cycling parameters were 3 min at 95°C, 34 cycles of 1 min at 95°C, 1 min at 51°C, 4 min at 72°C,
and finally 10 min at 72°C (parameter type 1). For the second reaction the primer pair was RNL-1 and RNL-5. Cycling parameters were 3 min at 95°C, 34 repeats of 1 min at 95°C, 1 min at 56°C, 4 min at 72°C and a final elongation of 5 min at 72°C (parameter type 2). The *Taq* polymerase from GE Healthcare (Otelfingen, Switzerland) included 2 mM $MgCl_2$, 0.5 µM of each primer and 0.25 mM of each desoxynucleotide in the master mix.

Based on these results and data from Raab *et al*. (2005), new primers with improved specificity were developed and applied in a nested PCR approach that was used for most other samples. Forward and reverse primers RNL-28a and RNL-5, respectively, were used in the first reaction and RNL-29 and RNL-30, respectively, in the second reaction (for primer sequences, see Table 2). *Taq* polymerase (master mix see above) or Phusion High-Fidelity DNA Polymerase from FINNZYMES (BioConcept, Allschwil, Switzerland) including 1x Phusion HF Buffer, 0.5 µM of each primer and 0.2 mM of each desoxynucleotide in the master mix, were used. Parameter type 1 (Table 1) was applied for both reactions when using *Taq* polymerase. Cycling parameters were changed when using Phusion polymerase: 30 s at 98°C, 33 cycles of 10 s at 98°C, 30 s at 55°C and 2 min at 72°C, followed by 10 min at 72°C (parameter type 3). Parameter type 4 (30 s at 98°C, 34 cycles of 10 s at 98°C, 30 s at 60°C, 2 min at 72°C, and 5 min at 72°C/Phusion polymerase) and parameter type 6 (see next paragraph) were used for exceptions (details summarized in Table 1). DNA was extracted at least twice from each isolate. If DNA was extracted from organic matter, results were confirmed by PCR products from spores in all isolates except JA202, NB102C and SW205.

Table 1 PCR conditions for isolates of *G. intraradices* and colonized plant roots (shaded in grey). Nested PCRs based on the same DNA extraction of one isolate are given in smaller case letters instead of numbers. Nested PCRs based on the same DNA extract like ITS sequences are shown in boldface. fwd (forward), iD (DNA of organic matter from inoculum), iS (spore from inoculum), nr (nested reaction), pD (DNA of organic matter from pot culture), rD (DNA of plant roots), rev (reverse), rS (spore from root organ culture).

Origin isolate or plant sample	**nested PCR**	**DNA template** origin/ dilution	**RNL PCR primers** 1. nr fwd	1. nr rev	2. nr fwd	2. nr rev	**Dilution** of 2. nr template	**parameter type** 1. nr	2. nr	**Further application**
AU212B	**1**	1xiS/no	28a	5	29	30	no	1	1	1x sequenced
	2	iD/1:10	28a	5	29	30	1:10	3	3	RFLP
CA502	**1**	1xiS/no	28a	5	29	30	no	1	1	1x sequenced
	2	iD/1:10	28a	5	29	30	1:100	3	3	RFLP
CC-4	**1**	>15rS/1:10	3	9	1	5	no	1	2	2x sequenced/RFLP
	2	15xrS/no	3	9	1	5	no	1	2	1x sequenced/RFLP
CR316A	1	1xiS/no	28a	5	29	30	no	1	1	1x sequenced
	2	iD/1:10	28a	5	29	30	no	3	3	RFLP
	3	iD/1:10	28a	5	29	30	no	1	1	RFLP
DAOM-197198	1	>15rS/no	3	9	1	5	1:100	1	2	5x sequenced
	2	>15rS/no	3	9	1	5	1:100	1	2	1x sequenced
DD-4	**1**	1xiS/no	28a	5	29	30	no	1	1	1x sequenced/RFLP
	2	pD/1:10	28a	5	29	30	no	1	1	RFLP
FL208A	1	iD/no	1	5	29	30	1:100	4	3	1x sequenced
	2	1xiS/no	1	5	29	30	no	4	3	RFLP
	3	1xiS/no	1	5	29	30	no	4	3	RFLP
	4	<15iS/no	1	5	29	30	no	4	3	RFLP
	5	>15iS/no	1	5	29	30	no	4	3	RFLP
JA202	**1**	iD/no	28a	5	29	30	no	1	1	1x sequenced
	2	iD/1:10	28a	5	29	30	no	3	3	RFLP
JJ141	**a**	5xrS/no	28a	5	28a	31	no	1	1	1x sequenced
	b	5xrS/no	28a	5	46	30	no	1	6	1x sequenced*
	c	5xrS/no	28a	5	46	5	no	1	6	1x sequenced*
JJ145	a	5xrS/no	28a	5	28a	31	no	1	1	1x sequenced
	b	5xrS/no	28a	5	46	30	no	1	6	1x sequenced*
	c	5xrS/no	28a	5	46	5	no	1	6	1x sequenced*
JJ183	1	4xrS/no	3	9	1	5	no	1	2	2x sequenced
	2	<15rS/no	3	9	1	5	no	1	2	1x sequenced/RFLP
KE114	**1**	1xiS/no	28a	5	29	30	no	1	1	1x sequenced
	2	iD/1:10	28a	5	29	30	no	3	3	RFLP
NB102C	**1**	iD/1:10	1	5	29	30	no	4	3	1x sequenced
	2	iD/1:10	28a	5	29	30	no	3	3	RFLP
SW205	1	iD/1:10	28a	5	29	30	no	3	3	RFLP
VA110	**1**	1xiS/no	28a	5	29	30	no	1	1	1x sequenced
	2	1xiS/no	28a	5	29	30	no	1	1	RFLP
	3	iD/1:10	28a	5	29	30	no	1	1	RFLP
Trifolium sp. S6	1	rD/1:100	28a	5	29	30	1:100	1	1	2x sequenced/RFLP
Trifolium sp. S10	1	rD/1:100	28a	5	29	30	1:100	1	1	2x sequenced/RFLP
Plantago lanceolata da2	1	rD/1:10	28a	5	29	30	no	1	1	2x sequenced/RFLP
Plantago lanceolata da4	1	rD/1:10	28a	5	29	30	no	1	1	1x sequenced/RFLP

* directly sequenced PCR product

Table 2 The sequences of RNL primers used to amplify regions of the mitochondrial rRNA large subunit gene (mtLSU) gene.

RNL primer	sequence	RNL primer	sequence
1*	AGACCCGAARCCWRGTGATCT	31	TTMGTGCCGCCACTTATTAG
2*	GGRAASAGCCCAGAAYA	33	CTGCCCTATAGAAGAGTTAC
2c	TCGTGATAAGGCGATTCGTC	35	TAACCCCTCAACGACCACAC
3*	TGCATMATGGGTCAGCGAGT	36	CACCTGTTCTGGGCTGTTGC
5*	GAGCTTCCTTTGCCATCCTA	36b	CACCAGTTCTGGGCTGTCGC
7*	CTTCTGCTTTCGGCGAAGAG	37	TAGCTGGGCTAAGAATGCTG
7b	CAGCTATGTCCACCGGCTCA	38	AGCTTGGACTAACCCACTAATG
9*	CAGTAAAGCTGCATAGGGTCT	39	CGGACCTATTGCCCAATACT
10*	AGAAGAAAGAGCTGGCTGTG	40	CCTAGAAGTGCTGGCGTGTT
10h	CCTAATAATCCTCCTACAAG	41	GGCTCTTTCCYGAACCTTAC
11*	AAGGCAACACGCCAGCACTT	42	TAGTACACTGCTAAGCTAGA
11b	AGGGCAACACGCCAGCACTT	43	CCCCTAACAGTTAATAACTC
12*	GATAGCGTAACAGCTCAGTG	44	CACTGCTAAGCTAGATTACT
12b	GATAGGGTAACAGCTCAGTG	45	ATACTCACCCTACCCTTACC
13*	TGGTCGATGGACGACGGATA	46	CCGACACAGGTCTGCAAGTA
13d	AGATCGMCGTAGTTCCTTCT	51	GATTCTGAAAGACGGAATAG
14*	AGGATAGGCCTGGAAACCAAGC	52	GGTTTTGCGTACGTAAGAGT
15*	CTGAGCTGTTACGCTATC	53	CATGGGTGTTGCTCTCAAATC
16*	ACCTGGAGATAGCTGGTCTT	54	AGCAACACCCATGTGCAAGT
17*	CCATAGAGTTGGCTCTAACA	55	ACTCTTACGTACGCAAAACC
17h	GCACGGAATTGAACCRTAGAG	57	CCTGTTAGGCGTACCTATGCC
24	GAGCATACTAAGGCGTAGAG	58	TAACCCATAATGCTCATCTA
25	ATCAGTGGGGAGAGGACAGT	59	GGAGATCATTGGGGTTATCC
26	CTTGTGCAAGTAGGCCTTCT	60	TAAGTAGGTGGACATAACTG
27	CCAACTATGCAACCGTAGG	61	GCCTAACCCTTCCCTATAGA
28a	CCATGGCCAAGTGCTATTTA	62	CCAGTGCCGTACCGTCTAGT
29	TAATAAGACTGAACGGGTGT	63	CATTATATGCTCCGGCGTAG
30	TAGCATCGGGCAGGTATCAG	64	AAGCACGGAATTGAACCATA

Shaded primers were used for the nested polymerase chain reactions; the others were used for sequencing.
* Primers designed by Raab *et al*. (2005).

6. PCR amplification of nuclear-encoded ITS rDNA

A nested PCR was performed according to Redecker (2000). The universal eukaryote primer pair NS5/ITS4 was used for the first PCR reaction (White *et al.* 1990). The *Glomus* group A-specific primer pair GLOM1310/ITS4i was used in the second PCR reaction. The *Taq* polymerase master mix contained following concentrations: 2 mM $MgCl_2$, 0.5 µM of each primer and 0.125 mM of each desoxynucleotide. Parameter type 5 cycling conditions for the first reaction were as follows: 3 min at 95°C, followed by 30 cycles of 30 s at 95°C, 30 s at 51°C, 2 min at 72°C, and a final extension phase of 10 min at 72°C. The second nested step was performed under the same conditions, but with an annealing temperature of 61°C (parameter type 6). Depending on success of PCR amplification, 1 µl PCR product of the first

reaction was used undiluted or at dilutions of 1:10 or 1:100 (in TE buffer) as template for the second reaction.

7. Cloning, sequencing and sequence analyses

The PCR products were purified using the High Pure Kit from Hoffmann LaRoche (Basel, Switzerland) and cloned into the pGEM-T vector (Promega/Catalys, Wallisellen, Switzerland) following the manufacturer's protocol. Before cloning, blunt-ended PCR products based on Phusion polymerase were incubated at 72°C for 13 min using *Taq* polymerase, 2 mM $MgCl_2$ and 0.125 mM dATP for adding 3′-adenines. Clones of mtLSU rDNA were amplified using the respective PCR primers of the second nested step or the vector primers M13fwd (GTA AAA CGA CGG CCA GTG) and M13rev (GGA AAC AGC TAT GAC CAT G). Products were purified and sequenced in both directions using the BigDye Terminator Cycle Sequencing Kit (ABI, Foster City, CA, USA) and an ABI 310 capillary sequencer. Sequencing primers for isolates differing from the sequencing set for JJ291 (Raab *et al*. 2005) are provided in Table 3. Complete sequences of the isolates JJ141 and JJ145 were composed by sequenced clones and directly sequenced PCR products (Tables 1 and 3). Clones of nuclear encoded ITS rDNA were sequenced in both directions using the primers of the second nested step or alternatively the universal forward primer ITS1F (CTT GGT CAT TTA GAG GAA GTA A) instead of GLOM1310. Sequences of mtLSU rDNA were aligned and corrected in BioEdit (Hall 1999), sequences of the ITS rDNA were edited in Sequence Navigator (version 1.0.1). Alignments were performed in BioEdit (Hall 1999) and in PAUP* 4.0b10 (Swofford 2001). DNA sequences were submitted to the European Molecular Biology Laboratory (EMBL) database under the accession numbers AM950203 to AM950227, and AM980833 to AM980863.

Isolates available as soil inoculum were extracted, amplified and analysed by RFLP or sequencing at least twice (Table 1).

8. Phylogenetic analyses

Phylogenetic trees were inferred using distance, parsimony or maximum likelihood criteria as implemented in PAUP*. Neighbor joining or heuristic search algorithms were applied for the respective criteria. Maximum likelihood models and parameters were

estimated using Modeltest 3.5 (Posada, 2004). In addition, Bayesian analyses were performed using MrBayes 3.1.1 for Macintosh (Ronquist & Huelsenbeck, 2003).

Insertions and deletions in the introns were coded by appending binary characters (1 for deletion, 2 for insertion) to the sequence matrix. Each deletion of more than three bases was coded, resulting in 21 binary characters added to the whole dataset. Regions of exons and introns that could not be aligned unambiguously were excluded from the analyses. Phylogenetic networks were obtained using SplitsTree 4.8 (Huson & Bryant, 2006). The Neighbor Net option using uncorrected distances and equal angles was chosen.

Table 3 Sequencing primer sets used for *Glomus intraradices* isolates and root colonizing *G. intraradices* (shaded in grey) which are different from the sequencing set for JJ291 (Raab et al., 2005). F (M13fwd), R (M13rev).

Origin isolate or plant sample	**Accession Number**	**Sequencing RNL-primers**	
		fwd	**rev**
CC-4	AM950204	1, 16, 2, 12b, 24, 25	5, 26, 27, 7b, 15, 14
	AM950205		
	AM950206		
CR316A	AM950207	F, 16, 35, 12, 37, 24, 25	R, 26, 27, 7, 7b, 38, 36
AU212B	AM950214		
JJ141	AM950208*	F, 11, 13d, 16, 35, 37, 46, 44, 53, 55	R, 33, 15, 14, 39, 43, 40, 30, 52, 54, 27, 5
JJ145	AM950209*		
JJ183	AM950210	1, 11, 13d, 16, 2, 12	5, 7b, 15, 14, 17, 10
	AM950211		
	AM950212		5, 7b, 15, 14, 17h, 10h
DD-4	AM950213	F, 11b, 2, 12, 24, 44, 25	R, 26, 27, 41, 15, 36, 39
JA202	AM950215	F, 16, 35, 12, 37	R, 33, 15, 36
KE114	AM950216	F, 16, 12, 37	R, 33, 15
NB102C	AM950217	F, 11, 13, 16, 35, 37, 24, 25	R, 26, 27, 7, 38, 36, 17, 10h
CA502	AM950218	F, 11, 16, 35, 12, 37, 24, 63	R, 59, 27, 7, 15, 62, 64
VA110	AM950219		
FL208A	AM950220	F, 58, 16, 60	R, 57, 15, 61
Plantago lanceolata da4	AM950221	F, 11b, 2c, 12, 24, 25	R, 26, 27, 7, 15, 36, 39
Trifolium sp. S6	AM950223	F, 11b, 16, 35, 37	R, 33, 15, 36b, 39
	AM950222		R, 33, 15, 36, 39
Trifolium sp. S10	AM950224		
	AM950225		
Plantago lanceolata da2	AM950226	F, 2c, 12, 24, 42, 25	R, 26, 27, 41, 15, 36
	AM950227	F, 11, 13d, 16, 12, 24, 42, 25	R, 26, 27, 41, 15, 45, 51

* sequence was composed from cloned and directly sequenced PCR products (see Table 1)

9. RFLP analyses

Based on virtual restriction patterns of the sequence data of the mtLSU rDNA, a RFLP system was established in order to distinguish different sequence types. For RFLP analyses, 20 U *Dra*III, 2.5 U *Bsa*JI from New England Biolabs (BioConcept, Allschwil, Switzerland) and 2.5 U *Hin*dIII from MBI Fermentas (LabForce, Nunningen, Switzerland) were used per sample, respectively. For each reaction 8 µl of the final nested PCR products were digested overnight at 37°C (*Bsa*JI samples at 60°C) in a total volume of 15 µl. For visualization, 1.5% agarose gels (1% SeaKem-0.5% NuSieve; Cambrex Bio Science, Rockland, ME) were loaded with the total volume of the digestion products and run at 100 V for 1 h. Fragment lengths were determined using Quantity One (version 4.1.0). The RFLP patterns were compared with the virtual patterns using a modified spreadsheet developed by Dickie *et al.* (2003).

IV. Results

1. Diversity of ITS sequences and mtLSU haplotypes in *G. intraradices* isolates

Sixteen isolates of *G. intraradices* originating from geographic locations on five continents were analysed. These isolates included two strains (JJ291 and BEG75) previously studied by Raab *et al.* (2005). In ITS-based phylogenies (Fig. 5), most isolates grouped into a clade that included sequences originating from a single spore of isolate JJ291. This group is designated as '*G. intraradices* main clade' because members have been used to genetically define the species in field studies (Hijri et al., 2006; Sýkorová et al., 2007a). Isolate FL208A from Florida (USA), showed a closer relationship to *G. proliferum*, clustering with sequences from the same isolate previously obtained by another group (P. Sudarshana *et al.*, unpublished), and with sequences from isolates VA110 and KS906 from Virginia and Kansas (USA), respectively. Isolate VA110 also was analysed in the present study, but the resulting sequences did not cluster with FL208A. Instead, it showed a close relationship with CA502 from California (USA). The cluster containing VA110 and CA502 showed a tendency to group outside the *G. intraradices* main clade in distance and parsimony analyses, but this was not supported by bootstrap values or in Bayesian analyses. Sequences AJ872051 and AJ872052 constitute a clade of environmental sequences (Hijri *et al.* 2006) which is a clearly separated sister group to the *G. intraradices* main clade. Generally, bootstrap and posterior probability values were relatively low, which may be caused by a very strict alignment, which

left only 300 bp for analyses. Nevertheless, the topology of the tree was highly consistent between analyses. Omitting the VA110/CA502 sequences raised the bootstrap value of the *G. intraradices* main clade, indicating that unresolved sister clades may deteriorate the support for the main clade. To confirm that ITS and mtLSU sequence data originate from the same fungal genotype, sequences were amplified from the same spore for several isolates (Table 1).

Among the 16 cultivated isolates, 12 mitochondrial haplotypes were distinguished (Table 4). An additional five haplotypes were identified in five root samples. The exon–intron structure of the gene region among the isolates is graphically depicted in Fig. 6. The length of the analysed region of the mtLSU varied between 1070 bp and 3935 bp among isolates because of the presence/absence of introns at three locations, and considerable length variation within introns. One isolate, KE114 from Kenya, did not contain any of the three introns. Isolate DAOM197198 from Canada, which is used in the genome sequencing project, grouped with isolate JJ291 from Switzerland in haplotype I. Two additional haplotypes were found in isolates JJ183, JJ141 and JJ145, which originated from the same field site as JJ291 in Switzerland. In addition to DAOM197198/JJ291, three pairs of isolates showed the same haplotypes, respectively (JJ141/JJ145, VA110/CA502, CC-4/CR316A).

Sequence polymorphism of the introns correlated with length polymorphism. In other words, variation in intron lengths mirrored differences in intron sequences (Fig. 6, Table 4).

In order to further confirm sequence homogeneity of the mtLSU rDNA within the same isolate (Raab *et al*. 2005), at least three cloned PCR products obtained from two different ROC plates were sequenced for each of the isolates CC-4, JJ183 and DAOM197198. The clones (see Tables 1 and 4) differed from the consensus sequence on average by 0.35% (CC-4) 0.24% (JJ183) and 0.28% (DAOM197198), which is within the range of the misincorporation error of *Taq* polymerase (Cline *et al.* 1996).

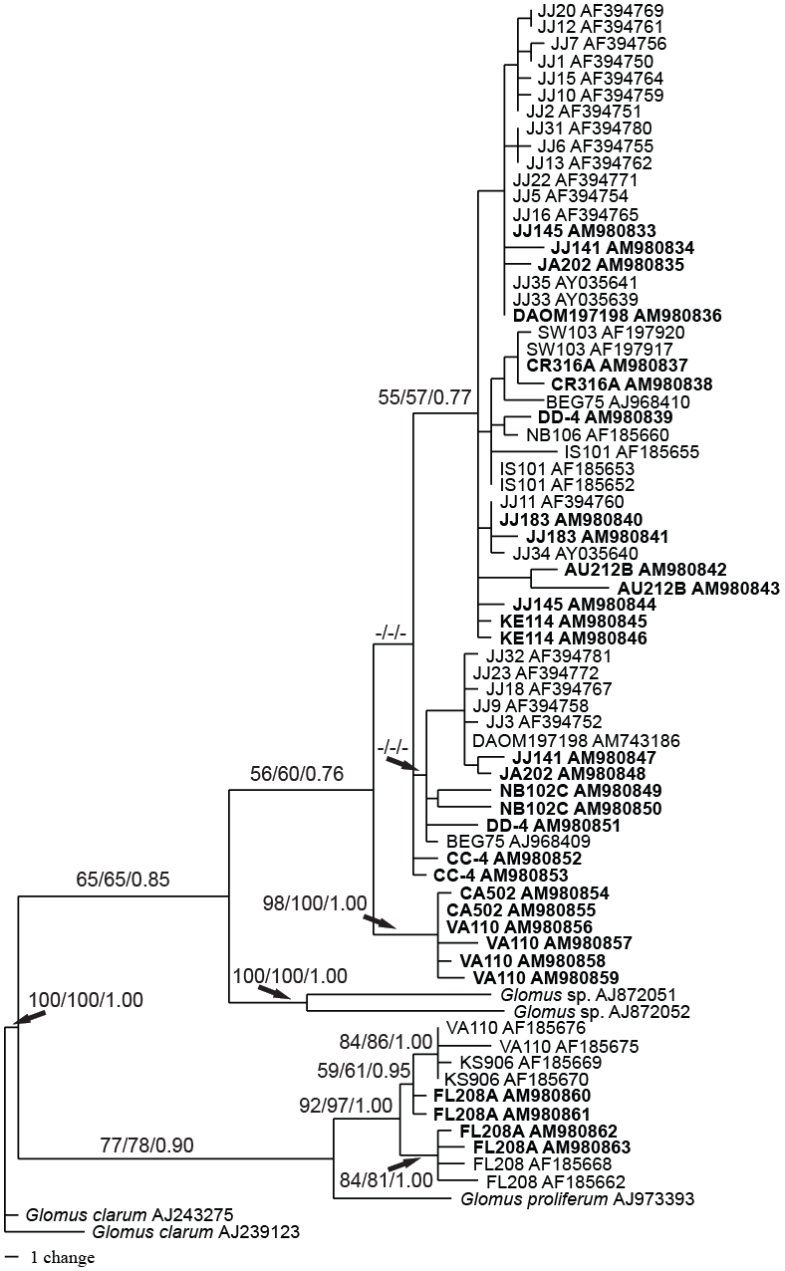


Fig. 5 Phylogenetic tree of *Glomus intraradices* isolates based on 5.8S rDNA and ITS2 sequences, with *Glomus clarum* as outgroup. The sequences JJ1 to JJ32 originated from a single-spore culture of *G. intraradices* JJ291 (Jansa et al., 2002b). The phylogenetic tree was generated following alignment of 300 characters and a heuristic search under the maximum parsimony criterion. Values at the nodes indicate: parsimony bootstrap values from 1000 replicates, neighbor-joining bootstrap values from 1000 replicates and Bayesian posterior probabilities. For clarity of the figure, only the values above 50% for the first five nodes from the root are provided. Sequences in bold type were unique to this study.

Table 4 Mitochondrial rRNA large subunit gene (mtLSU) rDNA sequence structure of *Glomus intraradices* isolates and root colonizing *G. intraradices* (shaded) within the priming sites of RNL-29/RNL-30.

Isolate/ plant sample (Origin)	**Clones** Accession number	**Haplotype**	**Introns**			**Exon region**	**Fragment length** (bp)
			Pos. 1 (-type, bp)	Pos. 2 (-type, bp)	Pos. 3 (-type, bp)	Complete/parts between intron positions (bp)	
DAOM-197198 (Canada)	AJ841804	I	1-1, 1057	2-1, 401	no	1073/363, 672, 38	2531
	AJ841808**		1-1, 1058	2-1, 401	no	1063/363, 670, 30	2522
	AM950203		1-1, 1057	2-1, 401	no	1071/363, 670, 38	2529
JJ291* (Switzerland)	AJ973189		1-1, 1056	2-1, 401	no	1071/363, 670, 38	2528
	AJ973190		1-1, 1056	2-1, 401	no	1071/363, 670, 38	2528
	AJ973192		1-1, 1057	2-1, 401	no	1071/363, 670, 38	2529
CC-4 (Canada)	AM950204	II	no	2-2, 415	3-1, 809	1071/363, 670, 38	2295
	AM950205		no	2-2, 415	3-1, 808	1071/363, 670, 38	2294
	AM950206		no	2-2, 415	3-1, 809	1071/363, 670, 38	2295
CR316A (Costa Rica)	AM950207		no	2-2, 415	3-1, 809	1071/363, 670, 38	2295
BEG75* (Switzerland)	AJ938171	III	no	2-2, 414	no	1071/363, 670, 38	1485
	AJ938173		no	2-2, 414	no	1071/363, 670, 38	1485
	AM040984		no	2-2, 414	no	1071/363, 670, 38	1485
JJ141 (Switzerland)	AM950208	IV	1-2, 1107	2-1, 401	3-2, 1356	1071/363, 670, 38	3935
JJ145 (Switzerland)	AM950209		1-2, 1107	2-1, 401	3-2, 1356	1070/363, 669, 38	3934
JJ183 (Switzerland)	AM950210	V	1-3, 1094	2-3, 389	no	1071/363, 670, 38	2554
	AM950211		1-3, 1094	2-3, 389	no	1071/363, 670, 38	2554
	AM950212		1-3, 1094	2-3, 389	no	1071/363, 670, 38	2554
DD-4 (Netherlands)	AM950213	VI	1-4, 425	2-4, 303	3-3, 944	1067/363, 666, 38	2739
AU212B (Australia)	AM950214	VII	no	2-5, 430	3-4, 850	1071/363, 670, 38	2351
JA202 (Japan)	AM950215	VIII	no	2-1, 401	no	1071/363, 670, 38	1472
KE114 (Kenya)	AM950216	IX	no	no	no	1070/363, 670, 37	1070
NB102C (Namibia)	AM950217	X	1-1, 1057	2-6, 339	3-1, 809	1071/363, 670, 38	3276
CA502 (California)	AM950218	XI	1-5, 489	2-7, 233	3-5, 667	1087/377, 672, 38	2476
VA110 (Virginia)	AM950219		1-5, 489	2-7, 233	3-5, 667	1088/377, 673, 38	2477
FL208A (Florida)	AM950220	XII	1-6, 662	no	no	1124/358, 728, 38	1786
Plantago lanceolata da4 (Hungary)	AM950221	XIII	1-7, 444	2-4, 303	3-6, 739	1067/363, 666, 38	2553
Trifolium sp. S6 (Switzerland)	AM950222	XIV	1-7, 444	2-4, 303	no	1067/363, 666, 38	1814
	AM950223		1-7, 444	2-4, 303	no	1067/363, 666, 38	1814
Trifolium sp. S10 (Switzerland)	AM950224		1-7, 445	2-4, 303	no	1067/363, 666, 38	1815
	AM950225		1-7, 444	2-4, 303	no	1067/363, 666, 38	1814
Plantago lanceolata da2 (Hungary)	AM950226	XV	no	2-4, 303	3-3, 944	1068/363, 667, 38	2315
	AM950227	XVI	1-3, 1094	no	3-7, 938	1067/363, 666, 38	3099
*Festuca pratensis** (Switzerland)	AJ841288	XVII	1-8, 1108	no	no	1069/366, 665, 38	2177
	AJ841289		1-8, 1108	no	no	1068/366, 664, 38	2176

Haplotypes and intron types were distinguished by sequence differences. Introns containing putative open reading frames for LAGLIDADG are shaded.
* Raab *et al* (2005); ** sequence incomplete at 3′-end.

2. MtLSU exon and intron phylogeny

The MtLSU exon sequences clearly separated the FL208A isolate from all other *G. intraradices* isolates in the phylogenetic tree (Fig. 7), a pattern that was in agreement with ITS phylogeny. Isolate FL208A grouped closer to *G. proliferum*. Isolates CA502/VA110 and two

sequences obtained from colonized roots of *Festuca pratensis* in a calcareous grassland (Raab et al., 2005) also grouped in clades with high bootstrap support. The exon and intron sequences of CA502 and VA110 did not differ by more substitutions than expected from *Taq* polymerase error.

All other *G. intraradices* isolates were quite similar in their mtLSU exon sequences (Fig. 7). A subclade containing mostly environmental sequences from grasslands received some bootstrap support in the exon tree (Fig. 7) and was also distinct in intron length (Fig. 6) and sequence (Table 4). Position 2 introns all were 303 bp in length and differed by only a few point mutations. All position 1 introns in this subclade were 425–444 bp long and lacked an open reading frame (ORF) for a homing endonuclease, which was detected in other isolates (see below).

Overall, sequences of the position 2 intron showed considerable similarity. This intron was present in 14 of 16 isolates and 6 of 9 environmental sequences. Phylogenetic analysis of the position 2 intron (Fig. 8) confirmed trends obtained from exon sequences. NeighborNet networks were used to provide more detailed visualization of any potential conflicts between intron phylogenies that might be caused by reticulate evolution. Four major groups of isolates were distinguished: the group of predominantly environmental sequences discussed above; a clade comprising all Tänikon isolates, DAOM197198 and JA202; a clade comprising BEG75, CC4 and CR316; and CA502/VA110. Some isolates did not fall into any of these groups, such as NB102C and AU212B. Isolate NB102C was not positioned on a distinct branch in the NeighborNet network in Fig. 8, possibly because the sequence region distinguishing the BEG75 and Tänikon isolate group was missing in this isolate.

A position 1 intron occurred in 10 of 16 isolates and 8 of 9 environmental sequences. A number of sequences containing this intron contained ORFs for homing endonucleases of the LAGLIDADG 2 type (Dalgaard et al., 1997). Furthermore, a LAGLIDADG type 1 ORF was found in isolates JJ141 and 145 in a position 3 intron. The position 2 intron did not contain any putative ORFs. Interestingly, some ORFs consisted of two regions separated by a putative noncoding sequence (Fig. 6). Phylogenetic analysis of position 1 intron sequences (Fig. 9) indicated that they were homologous. Some isolate groups described above were verified and the group comprising the Tänikon isolates and DAOM197198 was differentiated further into subgroups. Isolate NB102C associated closely with DAOM197198/JJ291, clarifying its ambiguous grouping in Fig. 8.

The position 3 intron was present in only 56% of the isolates. Although highly polymorphic in the central region, sequences of this intron were homologous in regions adjacent to the exons. In JJ141/JJ145, this intron contained an ORF for a type 1 homing endonuclease. Conflicts among exon and intron phylogenies were not detected, ruling out frequent transfer of these noncoding regions that could impair their use as intraspecific molecular markers.

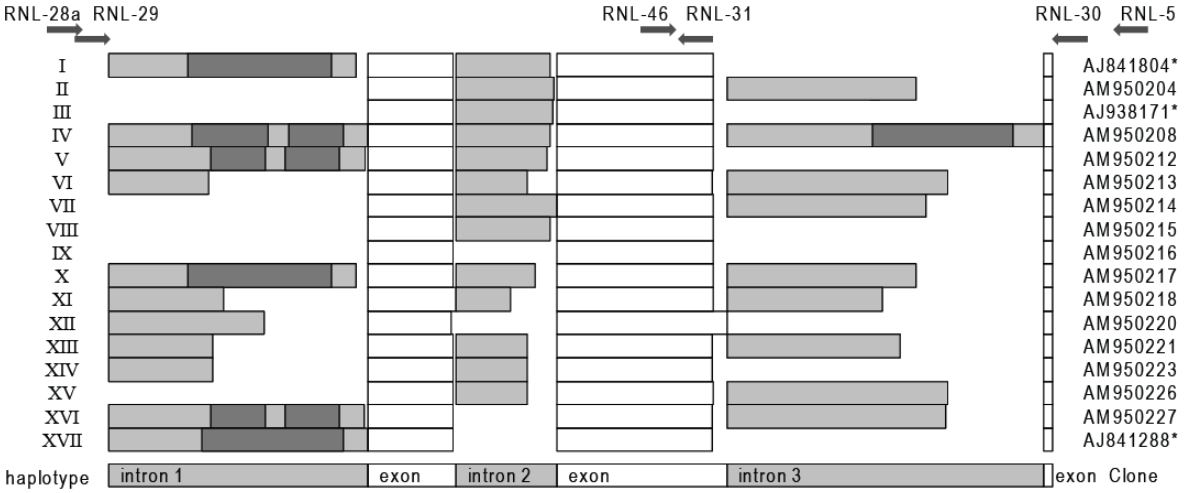


Fig. 6 Organization of the mitochondrial rRNA large subunit gene (mtLSU) region containing three exons and two to three introns for *Glomus intraradices* haplotypes I–XVII in 5′–3′ orientation. Introns are shaded in light grey and putative LAGLIDADG open reading frames (ORFs) in dark grey. Arrows show location and orientation of primers. *, Clones from Raab et al. (2005). Approximately to scale.

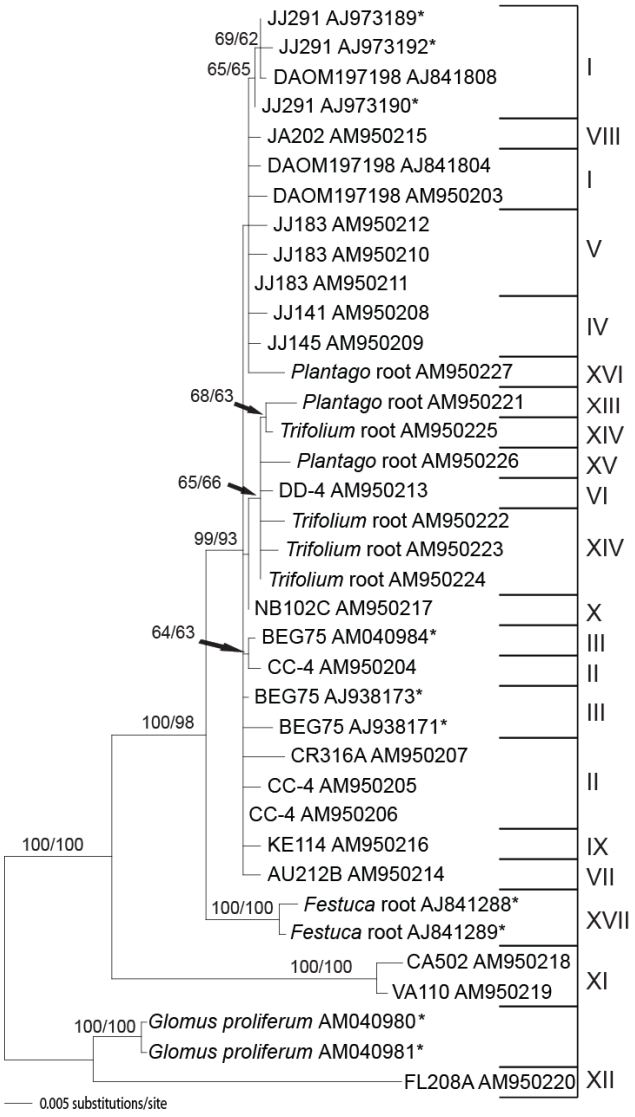


Fig. 7 Phylogeny of *Glomus intraradices* isolates and *Glomus proliferum* based on mitochondrial rRNA large subunit gene (mtLSU) exon sequences. The tree was rooted by midpoint rooting. Roman numerals indicate mtLSU haplotypes. The phylogenetic tree was obtained from 943 characters using a heuristic search under the maximum likelihood criterion. Values on the nodes indicate: neighbor-joining bootstrap values from 1000 replicates and maximum parsimony bootstrap values from 1000 replicates. Sequences from cultured isolates of *G. intraradices* are labeled with isolate codes and accession numbers. For sequences from roots, host species and accession number are indicated. *, Sequences from Raab et al. (2005).

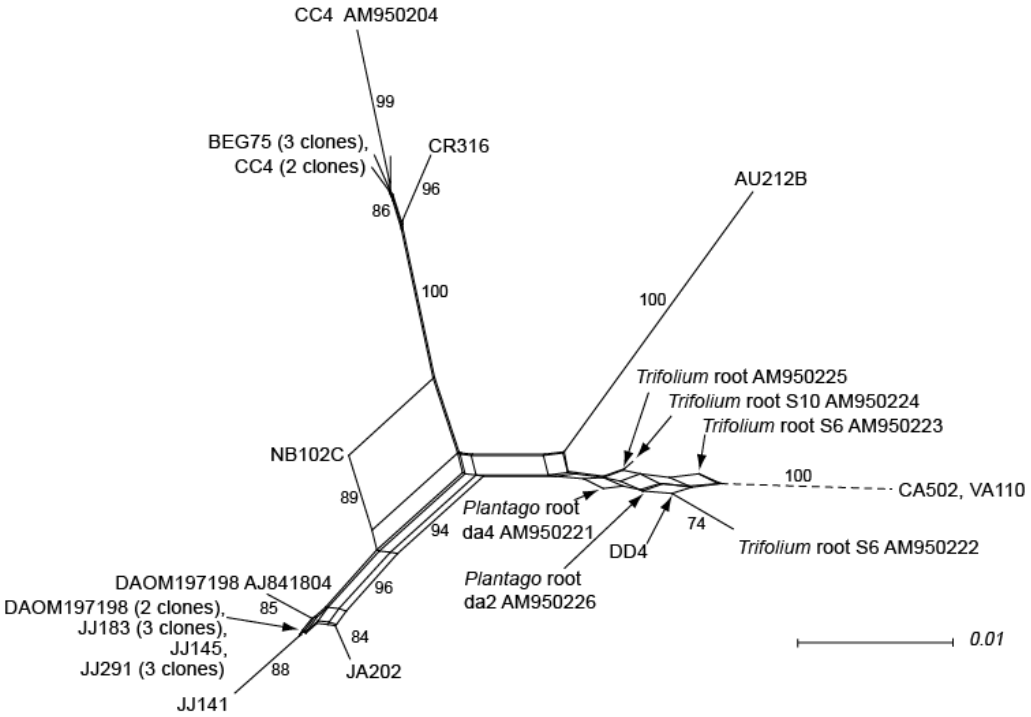


Fig. 8 NeighborNet network obtained from sequences of the position 2 intron. The dashed line indicates a branch that was reduced in length by a factor of 10 to improve readability of the figure. Numbers on the branches are bootstrap values from 1000 replications.

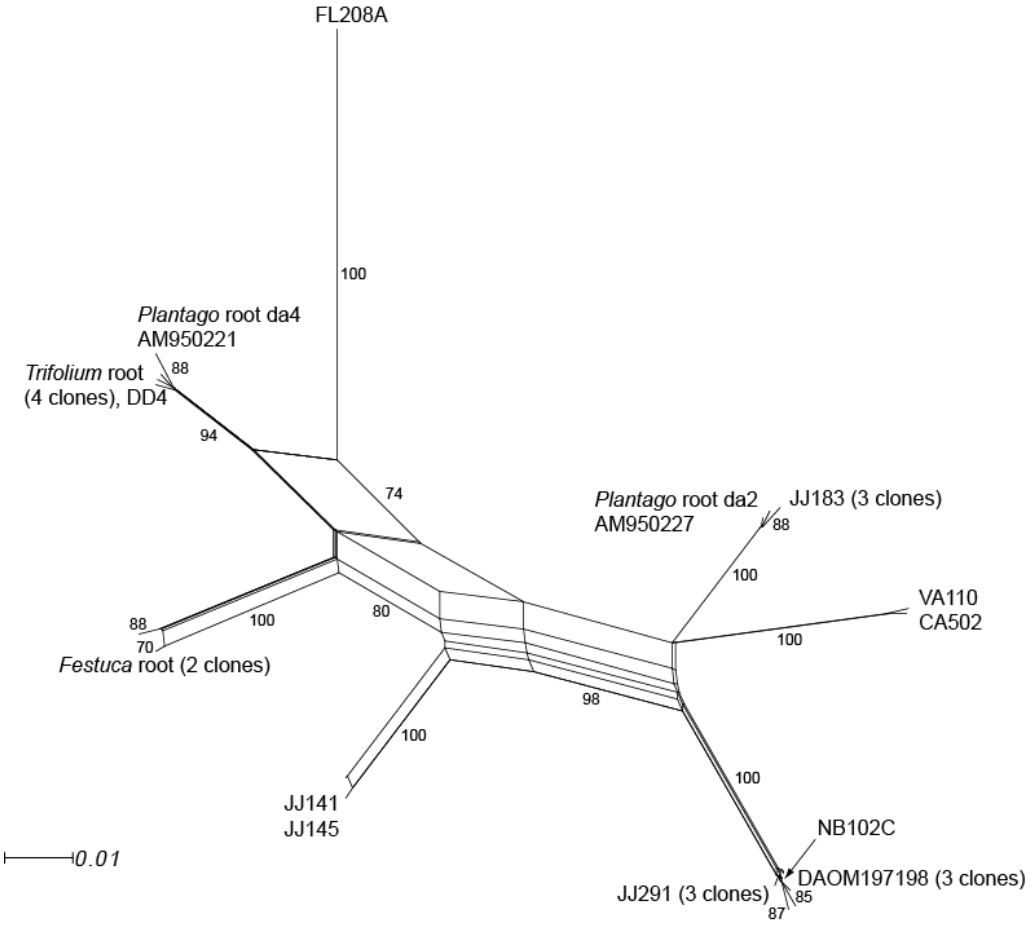


Fig. 9: NeighborNet network obtained from sequences of the position 1 intron. Numbers on the branches are bootstrap values from 1000 replications.

3. Intron stability in the mtLSU of *G. intraradices*

Comparisons of multiple culture lineages of the same isolates did not reveal any changes in the intron length or sequence of any isolate. For example, two lineages of isolate DAOM197198 were identical in both exon and intron sequences, even though one was obtained directly from G. Bécard (Toulouse, France) and has been propagated in the Botanical Institute in Basel since 1995 and the other was obtained from N. Requena (Karlsruhe, Germany) in 2005. The same result was obtained for two lineages of JJ291 cultivated independently over a 2-yr period. JJ141 and SW205 are two culture lineages originating from the same isolate and showed identical haplotypes. SW205 was pot cultured repeatedly at INVAM, whereas JJ141 was obtained as root organ culture. Identical haplotypes in populations from geographically distant locations (Switzerland/Canada, Virginia/California, Costa Rica/Canada) provided additional evidence for stability of markers.

4. Detection of haplotypes in field-collected roots

The mtLSU region could be amplified from field-collected roots using the improved primer combinations RNL-28a/RNL-5 and RNL-29/RNL-30. All amplified sequences clustered in the *G. intraradices* clade (Fig. 7). This result clearly demonstrated high primer specificity for the target clade. Interestingly, grassland isolate DD-4 was the only cultivated genotype clustering in the exon clade, which contained most environmental sequences from grassland communities. Among root samples, only one from *Plantago lanceolata* (sample code da2) yielded sequences of two different haplotypes (XV and XVI, Table 4).

5. RFLP analyses

From sequence data of the DNA region between primers RNL-29 and RNL-30, a combination of restriction enzyme sites were identified that unambiguously separated mtLSU haplotypes without sequencing (Table 5). A single conserved target site for *Dra*III was present in each mtLSU rDNA sequence. The restriction site was located in the exon region adjacent to the 5′-insertion site of position 2 intron. Restriction sites for *Bsa*JI were more abundantly distributed, but situated mainly nearer the 5′ ends of sequences. Restriction sites for *Hin*dIII are more abundant near the 3′ end of the DNA region.

Use of *Bsa*JI alone distinguished all haplotypes, whereas *Dra*III or *Hin*dIII clearly identified 59% and 71%, respectively, of these haplotypes. To avoid ambiguous identification as a result of similar restriction patterns, using all three enzymes provided the most accurate assessment of haplotypes.

Table 5 Restriction fragments lengths of mtLSU rDNA sequences calculated for sequenced clones from *G. intraradices* isolates and mycorrhizal roots, based on the target sites of the restriction enzymes *Dra*III, *Bsa*JI and *Hin*dIII. The ends of all complete fragment lengths are defined by the priming sites for RNL-29 and RNL-30. Different RFLP-subtypes ("Intra-types") of the same haplotype are indicated by lower case letters. Subtypes b and c were most likely caused by minor sequencing errors, as they were only found in clones, but never detected in original PCR products.

Clones Accession number	**Haplo-type**	**Intra-type**	**Com-plete frag-ment length** (bp)	***Dra*III** 5'...CACNNN▼GTG...3'	***Bsa*JI** 5'...C▼CNNGG...3'	***Hin*dIII** 5'...A▼AGCTT...3'
AJ841804	I	1	2571	1430, 1141	105, 1175, 1291	2069, 502
AJ841808**			2542	1431, 1111	105, 1176, 1261	2068, 474
AM950203			2569	1430, 1139	105, 1175, 1289	2067, 502
AJ973189*			2568	1429, 1139	105, 1174, 1289	2066, 502
AJ973190*			2568	1429, 1139	105, 1174, 1289	2066, 502
AJ973192*			2569	1430, 1139	105, 1175, 1289	2067, 502
AM950205	II	2a	2334	373, 1961	223, 438, 1673	1024, 508, 584, 218
AM950206			2335	373, 1962	223, 438, 1674	1024, 508, 585, 218
AM950204		2b	2335	373, 1962	44, 179, 438, 1674	1024, 508, 585, 218
AM950207		2c	2335	373, 1962	223, 438, 1674	1532, 585, 218
AJ938171*	III	3a	1525	373, 1152	223, 438, 864	1023, 502
AJ938173*			1525	373, 1152	223, 438, 864	1023, 502
AM040984*		3b	1525	373, 1152	44, 179, 438, 864	1023, 502
AM950208	IV	4	3975	1480, 2495	106, 1224, 2645	2117, 508, 459, 686, 205
AM950209			3974	1480, 2494	106, 1224, 2644	2116, 508, 459, 686, 205
AM950210	V	5a	2594	1467, 1127	106, 1211, 448, 829	2092, 502
AM950211			2594	1467, 1127	106, 1211, 448, 829	2092, 502
AM950212		5b	2594	1467, 1127	106, 1659, 829	2092, 502
AM950213	VI	6	2779	798, 1981	107, 541, 1224, 907	1337, 709, 733
AM950214	VII	7	2391	373, 2018	223, 2168	1039, 508, 626, 218
AM950215	VIII	8	1512	373, 1139	223, 1289	1010, 502
AM950216	IX	9	1110	373, 737	223, 887	609, 501
AM950217	X	10	3316	1430, 1886	105, 1175, 2036	2005, 508, 585, 218
AM950218	XI	11a	2516	877, 1639	106, 621, 1789	75, 1282, 498, 661
AM950219		11b	2517	877, 1640	106, 621, 1790	75, 1283, 1159
AM950220	XII	12	1826	1030, 796	105, 1721	433, 100, 710, 204, 379
AM950221	XIII	13	2593	817, 1776	107, 560, 1926	1356, 504, 733
AM950222	XIV	14	1854	817, 1037	107, 560, 1187	1356, 498
AM950223			1854	817, 1037	107, 560, 1187	1356, 498
AM950224			1855	818, 1037	107, 561, 1187	1357, 498
AM950225			1854	817, 1037	107, 560, 1187	1356, 498
AM950226	XV	15	2355	373, 1982	223, 1225, 907	913, 709, 733
AM950227	XVI	16	3139	1467, 1672	106, 1211, 921, 901	1703, 703, 733
AJ841288*	XVII	17	2217	1484, 733	106, 210, 1018, 883	609, 1111, 497
AJ841289*			2216	1484, 732	106, 210, 1018, 882	609, 1111, 496

* Raab *et al*. (2005), ** Sequence incomplete at 3'-end.

This RFLP approach was applied successfully to all isolates from INVAM, isolate DD-4 and all of the other sample types in this study (see Table 1, selected examples in Fig. 10). RFLP patterns were diagnostic for all of the haplotypes present in this study.

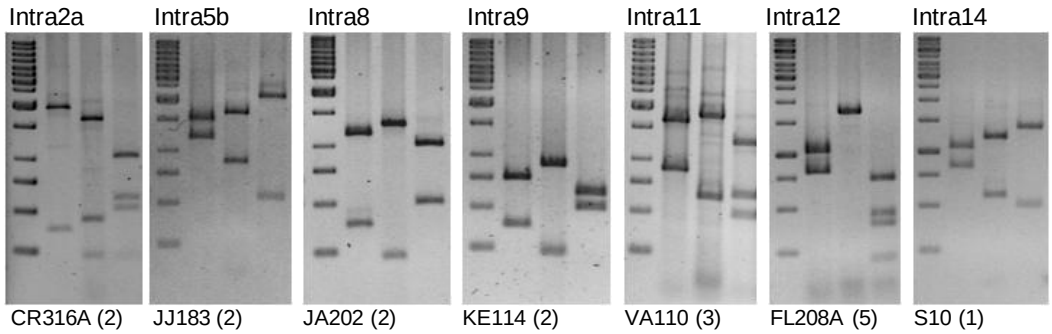


Fig. 10 Banding patterns of selected *Glomus intraradices* restriction fragment length polymorphism (RFLP) types. Respective isolate/root sample and polymerase chain reaction (PCR) product ID (in brackets) are shown below the pictures (for details see Table 1). The PCR products of the mitochondrial rRNA large subunit gene (mtLSU) rDNA generally were amplified using the primer pair RNL-29/30 in the second PCR step. RNL-1/5 was used for isolate JJ183, the only example of a cloned PCR product. DNA was digested with restriction enzymes *Dra*III, *Bsa*JI, *Hin*dIII and loaded in the same order onto gels. DNA ladder in left-most lane of each gel (bp): 250, 500, 750, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 5000, 6000, 8000, 10000.

V. Discussion

We showed in this study that intraspecific groups within *G. intraradices* were resolvable by variation and organization in a region of the mtLSU gene. Both exons and introns of this region provided stable molecular markers to identify haplotypes of this species from spores, colonized roots from field sites and even root fragments and mycelium mixed with organic matter obtained by sucrose extraction from pot cultures.

The AFLP markers were used initially to genetically differentiate intraspecific groups in *G. intraradices* (Koch et al., 2004), but they had limitations in that markers were nonspecific and could be analysed only using pure DNA from the target organism. DNA from microorganisms present associated with potcultured or field-collected AMF spores could confound AFLP results, so root organ cultures (Bécard & Fortin 1988) provided the means to obtain contamination-free DNA. *Glomus intraradices* was one of a limited range of AMF species compatible with this culture environment.

Recently, microsatellite simple sequence repeats (SSR) were developed for *G. intraradices* (Croll et al., 2008b; Mathimaran et al., 2008a). Five of the *G. intraradices* isolates resolved as five mtLSU haplotypes in this study were also resolved as distinct

genotypes by SSRs (Mathimaran et al., 2008a). The *G. intraradices* isolates from the Tänikon field site used in the present study allow the comparison with the multilocus genotypes of Croll *et al.* (2008), because some isolates were used in both studies. Based on the introns in positions 1 and 2, Croll et al. (2008b) distinguished the same three mitochondrial haplotypes (JJ291, JJ145/141 and JJ183). In addition, JJ145 and JJ141 could be distinguished by a single base pair length difference of one SSR locus. JJ291 and DAOM197198 were also differentiated by SSR data. These authors, like Raab et al. (2005), were not aware of the position 3 intron present in JJ141 and JJ145, which offers additional resolving power (e.g. among BEG75 and CC-4).

It is not yet clear if microsatellite SSR variation (Croll et al., 2008b) will elucidate the whole range of intraspecific diversity characterized by the mtLSU locus, because primers specific to *G. intraradices* genotypes were designed exclusively from isolates in root organ cultures, most of them originating from the Tänikon field site. Generally, multilocus population analyses from environmental samples face the problem that genotypes from different loci cannot be linked to each other if more than one target organism occurs in a sample. Mathimaran et al. (2008a) showed that SSR markers could be applied to measuring fungal diversity in colonized roots, but not yet under field conditions. Specificity must be tested exhaustively to exclude amplified products from plants and other associated microorganisms. Despite these potential problems, comparisons of both genotyping systems in more detail in future studies will offer new and intriguing insights.

Based on a greater breadth of isolate sampling, mtLSU data clearly show that intraspecific diversity within *G. intraradices* is considerably higher than previously reported. Croll et al. (2008b) reported that genotypic diversity in isolates from the Tänikon site was higher than that characterized in other isolates from Switzerland and one isolate from Canada (DAOM197198). Based on these comparisons, they concluded that intraspecific variation in *G. intraradices* is more diverse locally than globally, a hypothesis also put forward by Koch et al. (2004).

The finding that almost every *G. intraradices* isolate we sampled from a broad geographical range constitutes a different mtLSU haplotype certainly is surprising. Among 16 cultured fungal isolates, 75% comprised distinct haplotypes. Moreover, unique haplotypes were identified in most of the root samples analysed. Conversely, evidence also suggests that some haplotypes are distributed over a very broad geographical distribution, the most striking of them being haplotype I. Considerable effort, including sampling of multiple isolates from

each of a number of field sites around the world, will be needed to obtain a comprehensive overview of local versus pandemically distributed haplotypes in *G. intraradices*.

In contrast to SSR or AFLP fingerprinting methods, mtLSU sequences can be analysed phylogenetically, providing insights into the evolutionary relationships of the isolates and allowing to confirm the origin of the sequences from the target taxon.

Phylogenetic resolution within the *G. intraradices* clade was higher with mtLSU exon than ITS sequence data. The clade comprising sequences JJ1–JJ32 in the ITS trees has been used as a molecular grouping criterion for *G. intraradices* in field studies (Sýkorová et al., 2007a). Moreover, this clade also contains the isolate currently being sequenced to represent the genome of *G. intraradices* (Martin et al., 2004). MtLSU exon data clearly distinguish well-separated lineages within this group, among them the CA502/VA110 lineage, which is present in the ITS tree but its separation from the 'main clade' is not supported by bootstrap analysis.

A second clade of *G. intraradices* isolates can be distinguished in the ITS-5.8S phylogeny constructed in this study. It comprises isolates FL208, VA110 and KS906 grouping with *G. proliferum*. Some of these sequences have been present in the public database for years (e.g. AF185662, AF185668, AF185669, AF185670, AF185675 and AF185676 in Fig. 5). We confirmed that isolate FL208 groups in this clade, but even after repeated sequencing of multiple clones, ITS sequences from VA110 spores we obtained did not group with FL208A. Instead, its ITS and mtLSU sequences consistently grouped with CA502. The reason for this remains unclear. In any case, the mtLSU exon phylogeny confirms the genetic distance of FL208A from the other *G. intraradices* isolates. Based on unequivocal evidence from the ITS phylogeny, where a true outgroup could be used, the root of the tree is located between FL208/*G. proliferum* and the remaining taxa, which is consistent with the mtLSU exon phylogeny rooted by midpoint rooting (Fig. 7).

With its taxonomic resolution superior to ITS, the mtLSU exon region will be a useful molecular marker to contribute to a taxonomic consolidation of *G. intraradices.* Data from this and other gene regions will expand discovery of other isolates and clarify their interrelationships and distribution.

Fungal lineages generally are thought to lose introns more quickly rather than to gain them (Goddard & Burt, 1999). The complete absence of mtLSU introns in the KE114 isolate of *G. intraradice*s, and in *G. proliferum* may therefore represent a derived condition. The hypothesis that the introns were inserted as independent events in all other isolates is not

parsimonious. Introns in similar positions (1 and 2) have also been found in the more distantly related isolates of *G. mosseae* (Thiéry et al., 2010; see appendix). Possibly, these introns evolved in an ancestor common to both lineages. As in the study by Raab et al. (2005), ORFs coding for putative homing endonucleases were detected in some of the introns. Most of them were present in position 1 introns but an endonuclease ORF was also found in the position 3 intron. Homing endonucleases catalyse the spread of the intron-containing allele that encodes them to other intron-less alleles (Dalgaard et al., 1997).

Evidence has been accumulating that not all symbiotically active AMF necessarily sporulate in field settings (Hempel et al., 2007). The frequently used 'trap culturing' approach to propagate AMF from field samples in order to obtain spores for morphological analyses may bias the range of species-level taxa detected (Sýkorová et al., 2007a). A similar bias may be expected among isolates of a species within a population. Thus, it is highly desirable to develop specific molecular tools as culture-independent techniques to analyse intraspecific genetic diversity of glomeromycotan fungi directly within mycorrhizal roots. The possibility of amplifying and then characterizing glomeromycotan fungi that do not show evidence of sporulation in a field setting (Rosendahl & Stukenbrock, 2004) and therefore are not culturable is a significant asset of mtLSU primers. MtLSU markers targeting higher-level phylogenetic taxa can be developed to easily obtain sequence data from other species. By contrast, design of primers to characterize SSR loci in isolates of other species will require considerably more genomic information.

As the RFLP method to detect haplotype variation within *G. intraradices* was tested using a worldwide sampling of isolates, it can be expected to be applicable to the whole range of diversity found in this species. No laborious cloning steps are necessary to analyse PCR products, and direct sequencing is possible unless several haplotypes are present in the same root. In some cases sequencing may still be useful to obtain additional information about the haplotypes present or to confirm the RFLP results. Such an approach will facilitate sampling over a broader geographic range and more diverse habitats. Biogeographic patterns can be elucidated and the role of human intervention can be examined in such studies, and the hypothesis that *G. intraradices* is a true generalist can be tested. Moreover, it will be interesting to determine whether some of the mtLSU haplotypes or haplotype groups correspond to ecotypes and how these correlate with degree of culturability.

VI. Acknowledgements

This work was funded by a grant by the Swiss National Science Foundation to the senior author, which is gratefully acknowledged. The authors would like to acknowledge Thomas Boller and Andres Wiemken at the Botanical Institute for continuing support, Jan Jansa and Marcel van der Heijden, Natalia Requena and Guillaume Bécard for providing fungal isolates, Kurt Ineichen for cultivating *G. intraradices*, Zuzana Sýkorová and Istvan Parádi for root DNA extracts, and Zuzana Sýkorová, Pascal Bittel, Tobias Mentzel and the technical staff at Hebelstrasse for helpful discussions and making things work.

Appendix 2: Diversity of mitochondrial large subunit rDNA haplotypes of *Glomus intraradices* in two agricultural field experiments and two semi-natural grasslands

Published in Molecular Ecology [2010] **19**: 1497–1511

Boris Börstler, Odile Thiéry, Zuzana Sýkorová, Alfred Berner and Dirk Redecker

I. Abstract

Glomus intraradices, an arbuscular mycorrhizal fungus (AMF), is frequently found in a surprisingly wide range of ecosystems all over the world. It is used as model organism for AMF and its genome is being sequenced. Despite the ecological importance of AMF, little has been known about their population structure, because no adequate molecular markers have been available. In the present study we analyse for the first time the intraspecific genetic structure of an AMF directly from colonized roots in the field. A recently developed PCR-RFLP approach for the mitochondrial rRNA large subunit gene (mtLSU) of these obligate symbionts was used and complemented by sequencing and primers specific for a particularly frequent mtLSU haplotype. We analysed root samples from two agricultural field experiments in Switzerland and two semi-natural grasslands in France and Switzerland. RFLP type composition of *G. intraradices* (phylogroup GLOM A-1) differed strongly between agricultural and semi-natural sites and the *G. intraradices* populations of the two agricultural sites were significantly differentiated. RFLP type richness was higher in the agricultural sites compared with the grasslands. Detailed sequence analyses which resolved multiple sequence haplotypes within some RFLP types even revealed that there was no overlap of haplotypes among any of the study sites except between the two grasslands. Our results demonstrate a surprisingly high differentiation among semi-natural and agricultural field sites for *G. intraradices*. These findings will have major implications on our views of processes of adaptation and specialization in these plant/fungus associations.

Keywords: agriculture, genetic differentiation, *Glomus intraradices*, mutualism, mycorrhiza, population structure.

II. Introduction

Arbuscular mycorrhizal fungi (AMF) form a mutualistic symbiosis with the majority of land plants. AMF are placed in the Glomeromycota, a fungal phylum established by Schüßler et al. (2001b). In this symbiosis, the plant provides carbohydrates for the fungus, which in turn makes nutrients such as phosphate or nitrogen available to the phytobiont (Smith & Read, 2008). It has been shown that underground diversity and species composition of AMF affect plant diversity and plant community composition above ground (van der Heijden et al., 1998). One of the most widespread species of AMF is *Glomus intraradices*, which has been detected in almost all studies of AMF diversity in field settings.

A recent study by Stockinger et al. (2009) has highlighted the need to redefine this species based on molecular phylogeny and morphology, because the type strain from Florida (FL208), which has apparently not been detected in any environmental study so far, and the ubiquitous fungal clade appear to be distinct. Until this dilemma is resolved in nomenclature, we refer to this widely-known clade as "*G. intraradices* GLOM A-1", in concordance with our own field studies based on nuclear-encoded rRNA genes (e.g. Hijri et al., 2006; Sýkorová et al., 2007b). Furthermore, we use the term '*G. intraradices* sensu lato' to comprise all clades previously identified as *G. intraradices* based on spore morphology.

Glomus intraradices GLOM A-1 has been detected in an amazing variety of habitats, like mountain meadows (Börstler et al., 2006), calcareous grasslands (Sýkorová et al., 2007a), alpine meadows (Sýkorová et al., 2007b), high-input and low-input agricultural field sites (Hijri et al., 2006), and geothermal sites in Yellowstone (Appoloni et al., 2008). It was shown to be a generalist with regard to its host preferences (Sýkorová et al., 2007b) and its life history strategy (Sýkorová et al., 2007a). *Glomus intraradices* GLOM A-1 is the most frequently-used model organism in AMF research (e.g. (Helber & Requena, 2008), and the isolate DAOM197198 is currently the subject of the first genome sequencing project of an AMF species (Martin et al., 2008). For the isolate FACE494 the complete mitochondrial genome is available (Lee & Young, 2009).

In contrast to phytopathogenic fungi (e.g. Hovmøller et al., 2008) or ectomycorrhizal mutualists (e.g. Kretzer et al., 2005), the study of the diversity of AMF populations has only been possible since relatively recently when suitable molecular markers became available (Rosendahl, 2008). In particular, the polymorphism of nuclear-encoded rRNA genes within the organism and within spores (Sanders et al., 1995) had been an obstacle to resolve intraspecies genetic diversity. Stukenbrock & Rosendahl (2005a) conducted a pioneering

study on *Glomus* species, widespread in disturbed settings. The same markers (i.e. the nuclear encoded genes *GmFOX2* and *GmTOR2* and a region of nuclear large subunit rDNA) were applied in subsequent studies (Rosendahl & Matzen, 2008; Rosendahl *et al.*, 2009) of *G. mosseae* and its close relatives *G. geosporum* and *G. caledonium*, but in all of these studies spores of these fungi had to be used. The possible discrepancy between symbiotically active AMF and the presence of their propagules has been known for a long time from molecular studies of AMF species communities (e.g. Clapp et al., 1995), and a similar bias may be inevitable when using spores for population studies.

Genotyping methods used in recent studies of a field population of *G. intraradices* even required fungal biomass from root organ cultures (ROCs), which is the only currently known approach to obtain spores and mycelium free of contaminating microorganisms from these obligate biotrophs (Koch et al., 2004; Croll et al., 2008b). Only a few AMF species can be grown in this culture system, demonstrating a strong bias during the cultivation procedure, which may also exclude *G. intraradices* genotypes not well adapted to the artificial culture conditions. Koch *et al.* (2004) used amplified fragment length polymorphism (AFLP) to characterize isolates from tillage and no-tillage treatments of a field experiment in Tänikon (Switzerland). The same set of isolates was analysed by Croll *et al.* (2008b) using co-dominant multilocus markers. In this landmark study, the authors demonstrated considerable genetic structure within the population as well as genotype-level host preferences. By comparing the Tänikon isolates to a genetically similar isolate from Canada, and some isolates from other locations in Switzerland, the authors concluded that the global diversity of *G. intraradices* was low compared with local diversity. Another set of multilocus markers was presented by Mathimaran et al. (2008a). Importantly, Croll et al. (2009) demonstrated that some of the isolates they could differentiate genetically formed anastomoses (hyphal cross-bridges) and exchanged genetic markers *in vitro*. The application of these multilocus genetic markers in field-collected roots would be possible after developing nested PCR primers based on flanking sequences of the repeat motifs as suggested by Croll *et al.* (2008b). However, the flanking sequences are often polymorphic themselves (Mathimaran et al., 2008b) and even if primers are successfully designed, the principal problem remains how to correlate the different loci from environmental samples which may contain several genotypes.

Raab et al. (2005) and Börstler et al. (2008) demonstrated that the mitochondrial rRNA large subunit gene (mtLSU) and its introns can be used to distinguish isolates of *G. intraradices*. The mtLSU markers apparently have a resolution lower than multilocus markers (Croll et al., 2008b). Nevertheless, 12 haplotypes were distinguished among 16 isolates of

G. intraradices collected worldwide (Börstler et al., 2008). In contrast to the previously mentioned multilocus sequence markers, the mtLSU of *G. intraradices* can be amplified specifically from colonized root samples, offering a less biased view of fungal populations. Furthermore, phylogenetic analysis of the mtLSU exon sequences allows to test whether the organism detected belongs to the target clade. Börstler *et al.* (2008) developed a RFLP approach that simplifies the identification of mtLSU haplotypes, allowing larger numbers of samples to be analysed by avoiding the laborious task of screening libraries of cloned sequences and reducing the amount of sequencing required.

On the species level, it has been shown that agricultural practice alters AMF community composition in the soil. Many arable soils, in particular in monocropping systems exhibit a reduced AMF diversity compared with natural ecosystems (Helgason et al., 1998a; Oehl et al., 2003; Hijri et al., 2006). Nevertheless, it is still not well understood how AMF communities adapt to environmental conditions, in particular due to the striking ubiquity of some species, especially *G. intraradices* GLOM A-1.

Within morphospecies, differences in the production of external hyphae and hyphal phosphorus uptake of *G. mosseae* isolates were reported by (Munkvold et al., 2004). Distinct growth parameters and effects on plant growth of different isolates of *G. intraradices* originating from Tänikon (see above) were observed by Koch et al. (2006). These findings suggest considerable functional differentiation within species, implying the need to study the genetic foundation of these differences and the spatial and temporal dynamics of population diversity.

In this study, we used the mtLSU PCR-RFLP-sequencing approach developed by Börstler et al. (2008) to investigate intraspecific genetic diversity of *G. intraradices* GLOM A-1 in field-collected roots. The efficiency of the RFLP approach was verified using a PCR primer set specifically designed for the haplotype I of *G. intraradices*, which comprises most isolates from Tänikon (Croll et al., 2008b), as well the isolate DAOM197198 from Canada, which is used for genome sequencing.

By analysing the mtLSU haplotype diversity of *G. intraradices* GLOM A-1 in two arable and two grassland field sites in Switzerland and France, we wanted to address the following questions:

- is population-level diversity decreased in arable soils?
- does *G. intraradices* show genetic differentiation between field sites or between different types of agricultural management?

- may the surprising ubiquity of *G. intraradices* be due to previously unresolved intraspecific groups, for example, ecotypes, as preliminary results on a 'grassland clade' had suggested (Börstler et al., 2008)?

The resulting data can be expected to provide a better understanding of how genetic diversity is partitioned in this ecologically important phylum of mutualists, in particular the model species *G. intraradices*, and better hints toward the processes by which this diversity is generated and maintained.

III. Materials and methods

1. Field sites and sampling

The agricultural study sites were the long-term tillage experiment in Tänikon (Switzerland; see Fig. 1), which was established in 1987, and a field experiment in Frick (Switzerland; see Fig. 2), which was established in 2002. For soil properties and other details in Tänikon see Jansa et al. (2002a) and Jansa et al. (2003), for the details in Frick see Berner et al. (2008). Both experiments included different treatment plots (six treatments in Tänikon, eight treatments in Frick) organized in a randomized block design in Tänikon or a split plot design in Frick. In both field experiments, all different treatments were replicated four times. Both sites were subject to crop rotation, from which *Zea mays* was sampled. Samples in Tänikon were taken in June 2007 and sampling in Frick was performed in July 2008. In Tänikon, complete root systems of a total of 48 plants were removed using a spade from the following treatment plots: no-tillage, chisel (loosening soil with a wing share chisel to the depth of 25 cm without turning soil upside-down), tillage (ploughed to the depth of 25 cm). In Frick, a total of 32 root cores were taken adjacent to the plants using a soil corer (diameter of 4 cm) to a depth of 10 cm. The treatment plots were: chisel [chisel plough (15 cm) including a stubble cleaner (5 cm)] and conventional tillage (mouldboard plough, operating at 15-cm depth). Crop residues were not removed in the sampled plots of Tänikon, in which mineral fertilization was applied. The sampled plots in Frick were fertilized with slurry, additional preparations were not applied. A similar sampling design was used in both experiments: four samples were taken at equal distance to each other along a transect crossing the middle of each treatment plot (19 m transects of 19 x 6 m plots in Tänikon, 17 m diagonal transects of 12 x 12 m plots in Frick), resulting in 16 samples per treatment. Samples covered with soil

were stored at 4°C immediately after transport. Within the next 4 days, the root systems of each plant sample were rinsed in tap water, fine root pieces were randomly placed in Petri dishes and washed again in distilled water. Finally, aliquots of 50–80 mg (fresh weight) root fragments per sample were frozen in liquid nitrogen and stored at −80°C.

The semi-natural sites were one calcareous grassland next to the Landskron castle (near the village Leymen, Alsace, France; see Fig. 3), and two adjacent sample areas from the alpine region near the village of Ramosch, Engadin region, Switzerland (see Fig. 4). Both semi-natural study sites were mown regularly and were species-rich grasslands. However, plant communities and soil properties differed between Landskron and Ramosch, for details see Sýkorová et al. (2007a,b), respectively. Both study sites were previously investigated with regard to their AMF diversity using ITS regions as marker genes by the latter authors. In these studies, sampling was performed in July 2002 and July 2005 in Landskron, in May 2003 in Ramosch. All semi-natural plant samples were removed randomly in soil cores due to a random distribution of different plant species. Sizes of the collection areas were approximately 15 m^2 in Landskron and approximately 40 m^2 and 30 m^2 in Ramosch. Roots were treated using the same procedures as described for the agricultural study sites. In this study, we have chosen DNA extracts that were screened positive for *G. intraradices* GLOM A-1 by Sýkorová et al. (2007a,b). In Landskron, *G. intraradices* was detected in the roots of 21 plant samples. Additionally, one *Medicago sativa* root sample originating from a compartment system established with inoculum from Landskron (Sýkorová et al., 2007a) was analysed for comparison. In Ramosch, 26 root samples were tested positive for *G. intraradices*, two samples were already analysed for mtLSU haplotypes by Börstler *et al.* (2008). The samples from the two Ramosch sampling areas were pooled as Sýkorová *et al.* (2007b) could not observe significant differences of AMF communities between both collecting areas. This approach was confirmed by initial analyses of *G. intraradices* mtLSU haplotypes in the present study. Plant species of the grassland samples are listed in Table 1.

Table 1 Plant species and samples tested positive for *Glomus intraradices* GLOM A-1 by Sýkorová *et al.* (2007a,b) in the grasslands of Landskron and Ramosch, respectively.

Landskron		**Ramosch**	
Plant species	**Number of root samples**	**Plant species**	**Number of root samples**
Bromus erectus	5	*Gentiana acaulis*	3
*Festuca pratensis**	3	*Gentiana verna*	4
Inula salicina	3	*Hieracium hoppeanum*	2
Medicago sativa	5	*Leontodon hispidus*	2
Origanum vulgare	5	Poaceae sp.	2
		Ranunculus montanus	1
		Thymus pulegioides	1
		Trifolium sp.**	11

* not published before, ** two samples previously analyzed for mtLSU haplotypes by Börstler *et al.* (2008).

An isolate of *G. clarum* (ID code MUCL46238), originating from La Palma (Pinar del Rio, Cuba), was cultivated as ROC and used as outgroup for phylogenetic analyses in this study. For DNA extraction, the medium was dissolved in 10 mM sodium acetate–citrate buffer (pH 6.0) and spores were washed in sterile water according to Doner & Bécard (1991).

2. DNA extraction

DNA extractions of all plant samples were performed using the DNeasy plant mini kit (Qiagen, Hilden, Germany): the complete root aliquots were ground in liquid nitrogen, followed by DNA extraction according to the manufacturer's instructions. Finally, DNA was eluted in two steps using 50 µl each. DNA extraction of spores from *G. clarum* was performed as described in Redecker et al. (1997).

3. PCR amplification of mtLSU

The nested PCR described in Börstler *et al.* (2008) was optimized for amplification from plant root DNA extracts on the Eppendorf Mastercycler epgradient S (Vaudaux-Eppendorf, Schönenbuch, Switzerland). For the first reaction the primers were RNL-28a and RNL-5. Phusion High-Fidelity DNA Polymerase from FINNZYMES (BioConcept, Allschwil, Switzerland) including 1 µl genomic DNA, 1x Phusion HF Buffer, 0.5 µM of each primer, 0.2 mM of dNTPs, 4x BSA and 3% DMSO in a volume of 25 µl were used. Cycling parameters were 30 s at 98°C, 37 cycles of 10 s at 98°C, 20 s at 57°C and 1 min 40 s at 72°C, followed by 10 min at 72°C. Products of the first PCR were diluted 1:100 in distilled water and used as

template for the second reaction consisting of the primer pair RNL-29 and RNL- 30. Phusion polymerase including 1 µl of diluted PCR product from the first reaction, 1x Phusion HF Buffer, 0.4 µM of each primer, 0.15 mM of dNTPs, 4x BSA and 3% DMSO in a volume of 50 µl were used. Cycling parameters were 30 s at 98°C, 36 cycles of 10 s at 98°C, 20 s at 60°C and 1 min 40 s at 72°C, followed by 10 min at 72°C. These improved PCR conditions were additionally tested on different isolates of *G. intraradices* used in Börstler *et al.* (2008), with PCR products ranging from 1110 to 3975 bp in length. Examples for PCR products from isolates and field collected roots are shown in Fig. 5. Reproducibility of the nested PCR approach was confirmed by testing all Tänikon samples twice. For samples from all study sites that revealed no product, PCRs were repeated at least once using positive controls. Besides samples already analysed by Börstler *et al.* (2008), mtLSU products of four further root samples from Landskron and the compartment system were amplified using *Taq* polymerase as described by Börstler et al. (2008). The mtLSU of *G. clarum* was amplified using the primers RNL-1 and RNL-5 (Raab et al. 2005). The PCR master mix was the same as described for the first nested reaction used for the root samples (see above). Cycling parameters were 30 s at 98°C, 37 cycles of 10 s at 98°C, 20 s at 58°C and 2 min at 72°C, followed by 10 min at 72°C.

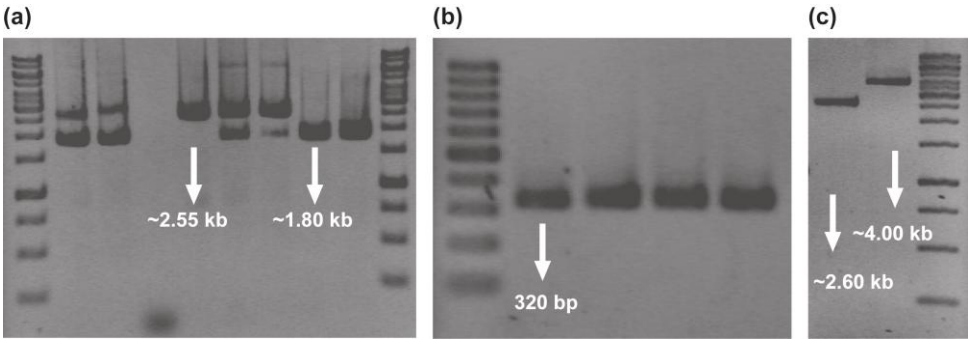


Fig. 5 (a) MtLSU PCR products from colonized root samples (Landskron) of the general PCR approach for *Glomus intraradices* (Börstler *et al.* 2008), **(b)** mtLSU PCR products from colonized root samples (Tänikon) of the specific approach for haplotype I, and **(c)** mtLSU PCR products of the isolates JJ183 (left lane) and JJ141 (middle lane) amplified with the adapted PCR conditions of the nested PCR approach developed by Börstler *et al.* (2008) for plant root DNA. For the general mtLSU PCR approach a 1 kb ladder (#SM0311) was used, for the haplotype specific approach a 100 bp ladder (#SM0242) was used (Fermentas Life Science; www.fermentas.com).

4. Specific PCR amplification of mtLSU haplotype I

A specific nested PCR approach was developed for the detection of mtLSU haplotype I of *G. intraradices*. For the first reaction the forward primer RNL-47 (GTT GAG GGG TGA

CCT TCA AT) and the reverse primer RNL-49 (GCT ACC TAT GCC GGG TTT TC) were designed. *Taq* polymerase from GE Healthcare (Otelfingen, Switzerland) was used with 2 mM $MgCl_2$, 0.5 µM of each primer and 0.25 mM of each desoxynucleotide in a volume of 25 µl per sample. Cycling parameters were 3 min at 95°C, 32 cycles of 30 s at 95°C, 30 s at 64.5°C and 2 min 30 s at 72°C, followed by 10 min at 72°C. Products of the first PCR were diluted 1:100 in distilled water and used as template for the next step. For the second reaction, the forward primer RNL-48 (AGC CTA AGG TCT TAG AGA CT) was newly designed. The primer RNL-10 (Raab *et al.* 2005) was chosen as reverse primer. The *Taq* polymerase master mix contained following concentrations: 2 mM $MgCl_2$, 0.5 µM of each primer and 0.125 mM of each desoxynucleotide. Cycling parameters were 3 min at 95°C, 30 cycles of 30 s at 95°C, 30 s at 64.5°C and 2 min at 72°C, followed by 10 min at 72°C.

5. RFLP analyses and definition of RFLP types

In order to distinguish different haplotypes, all mtLSU PCR products were analysed using the RFLP approach developed and described in detail by Börstler *et al.* (2008). For each sample, digestion products of the restriction enzymes were always loaded in the same order (*Dra*III, *Bsa*JI and *Hin*dIII) onto the agarose gels. The combined pattern of the three digestion products was defined as RFLP type 'Intra type' when sequence analyses (see below) of the respective PCR product matched with *G. intraradices* GLOM A-1. Different Intra types were distinguished by Arabic numerals corresponding to the Roman numeral of the mtLSU haplotype determined by complete sequence data (see next section), either already defined by Börstler et al. (2008) or in the present study (e.g. Intra8 of haplotype VIII). Occasionally, the same Intra type can comprise several mtLSU haplotypes which are not distinguishable with the RFLP approach. In that case the Arabic numeral corresponds to the Roman numeral of the mtLSU haplotype detected first (e.g. Intra3 of the haplotypes III and XIX). Different RFLP subtypes of the same mtLSU haplotype, indicated by lower case letters, were defined by Börstler et al. (2008). They were most probably caused by *Taq* polymerase errors and did not occur in this study using Phusion polymerase (Table 2). Different RFLP types of non-target taxa were distinguished by Arabic numerals.

Table 2 Restriction fragment lengths of mtLSU rDNA sequences calculated *in silico* for sequenced clones (except for the directly-sequenced product labeled by *) from root colonizing *Glomus intraradices* and the isolate MUCL46238 of *Glomus clarum*, based on the target sites of the restriction enzymes *Dra*III, *Bsa*JI and *Hin*dIII. The ends of all complete fragment lengths are defined by the priming sites for RNL-29 and RNL-30. RFLP types from *G. intraradices* are indicated as "Intra types". They were defined after previously detected Intra types by Börstler *et al*. (2008) or, if detected for the first time, after the number of the corresponding haplotype. A RFLP type from *G. clarum* ("Clar type") is added for comparison.

Clones/PCR product (Accession Number)	**RFLP type**	**Complete fragment length** (bp)	***Dra*III** Restriction site: 5'...CACNNN▼GTG...3'	***Bsa*JI** Restriction site: 5'...C▼CNNGG...3'	***Hin*dIII** Restriction site: 5'...A▼AGCTT...3'
FN377581	Intra1	2569	1430, 1139	105, 1175, 1289	2067, 502
FN377583		2569	1430, 1139	105, 1175, 1289	2067, 502
FN377580	Intra2a	2302	373, 1929	223, 448, 1631	998, 508, 578, 218
FN377578	Intra3a	1500	373, 1127	223, 448, 829	998, 502
FN377587		1525	373, 1152	223, 438, 864	1023, 502
FN377582	Intra5a	2594	1467, 1127	106, 1211, 448, 829	2092, 502
FN377576	Intra6	2807	826, 1981	107, 569, 1224, 907	1365, 709, 733
FN377589		2801	820, 1981	107, 563, 1224, 907	1359, 709, 733
FN377592		2798	817, 1981	107, 560, 1224, 907	1356, 709, 733
FN377577	Intra8	1512	373, 1139	223, 1289	1010, 502
FN377593	Intra13	2596	820, 1776	107, 563, 1926	1359, 504, 733
FN377597		2581	807, 1774	107, 550, 1924	1346, 502, 733
FN377598		2572	798, 1774	107, 541, 1924	1337, 502, 733
FN377599		2581	807, 1774	107, 550, 1924	1346, 502, 733
FN377600		2596	820, 1776	107, 563, 1926	1359, 504, 733
FN377591	Intra14	1854	817, 1037	107, 560, 1187	1356, 498
FN377594		1857	820, 1037	107, 563, 1187	1359, 498
FN377595		1848	811, 1037	107, 554, 1187	1350, 498
FN377596		1848	811, 1037	107, 554, 1187	1350, 498
AM950222**		1854	817, 1037	107, 560, 1187	1356, 498
AM950223**		1854	817, 1037	107, 560, 1187	1356, 498
AM950224**		1855	818, 1037	107, 561, 1187	1357, 498
AM950225**		1854	817, 1037	107, 560, 1187	1356, 498
FN377579	Intra20	2314	373, 1941	223, 2091	1010, 508, 578, 218
FN377584	Intra22	1904	1171, 733	103, 30, 11, 210, 667, 883	1407, 497
FN377585	Intra23	2713	1171, 1542	103, 30, 11, 210, 667, 1692	1407, 503, 585, 218
FN377586	Intra24	2259	373, 1886	223, 2036	948, 508, 585, 218
FN377588*	Intra25	3293	1171, 817, 1305	103, 30, 11, 210, 667, 2272	1407, 1631, 32, 223
FN377590	Intra27	3210	817, 2393	107, 560, 2543	1356, 504, 1145, 205
FN377601	Clar1	2217	1427, 790	106, 1171, 940	1677, 540

** Börstler *et al.* (2008)

6. Sequencing, cloning and sequence analyses

In addition to the RFLP approach, sequencing and cloning reactions were performed for different purposes: (i) confirmation that PCR products originated from *G. intraradices* GLOM A-1, (ii) facilitation of banding pattern interpretation in mixed samples, and (iii) increased resolution of RFLP patterns shared between different sites.

In order to confirm the identity of *G. intraradices* for every unknown RFLP pattern, exon regions from mtLSU PCR products were sequenced. At least one representative sample of each RFLP type of *G. intraradices* (Intra type) in every study site was sequenced completely for the detailed determination of the respective sequence type, which was defined as mtLSU haplotype according to Börstler *et al.* (2008). Based on these results, precise restriction fragment sizes were calculated *in silico* for each Intra type and are listed in Table 2. PCR products containing more than one mtLSU haplotype yielded mixed RFLP patterns. Therefore, these products were cloned in order to separate them for RFLP analyses and/or sequence analyses. For all single products, cloning was not necessary before sequencing, but it was also performed in order to receive enough template for complete sequencing reactions of mtLSU PCR products and to store clones as reference material. Furthermore, RFLP types shared by more than one study site were analysed at least in diagnostic sequence regions (see section 3.4) in order to see, if the same RFLP type is also represented by the same sequence type (see section 3.3.5). Additionally, all PCR products of the haplotype I-specific approach were sequenced.

For cloning and sequencing reactions, PCR products were purified using the High Pure Kit from Hoffmann LaRoche (Basel, Switzerland). Cloning was performed using the pGEM-T vector (Promega/Catalys, Wallisellen, Switzerland) as described in the manufacturer's instructions. Clones were amplified using the respective PCR primers of the second nested PCR reactions or the vector primers M13fwd (GTA AAA CGA CGG CCA GTG) and M13rev (GGA AAC AGC TAT GAC CAT G). For sequencing of PCR products or amplified clones the BigDye Terminator Cycle Sequencing Kit (ABI, Foster City, USA) and either the ABI 310 capillary sequencer or the AB3130xl genetic analyser were used. Sequencing primers for the respective samples are listed in Table 3. Sequences of the primers not published in Raab *et al.* (2005) and Börstler *et al.* (2008) are provided in Table 4. Sequences were initially edited in Sequence Navigator (version 1.0.1) and corrected and aligned in BioEdit (Hall 1999). Complete sequences were submitted to the European Molecular Biology Laboratory (EMBL) database under the accession numbers FN377576–FN377601.

Table 3 Sequencing primer sets used for completely sequenced mtLSU PCR products of *Glomus intraradices* and sequencing primers used for diagnostic sequence regions of PCR products corresponding to RFLP types of *G. intraradices* (Intra types) shared by more than one study site (shaded). Additionally, sequencing primers used for the determination of RFLP types belonging to other fungi than *G. intraradices* ("Fung types") are given. F (M13fwd), R (M13rev).

RFLP type	Haplotypes of completely sequenced Intra types	Forward sequencing RNL-primers	Reverse sequencing RNL-primers
Intra1	I	F, 29, 11, 13, 16, 2, 12, 37	R, 30, 7, 7b, 15, 14, 17, 10
Intra2a	XXI	F, 16, 35, 12, 37, 24, 25	R, 26, 27, 7, 7b, 38, 36
Intra3a	XIX	F, 16, 2, 35, 12, 37	R, 33, 15, 36
	III	F, 16, 2, 35, 12, 37	R, 33, 7, 15, 36
Intra5a	V	F, 11, 13d, 16, 2, 12, 37	R, 30, 7, 15, 14, 17, 10
Intra6	XVIII	F, 11b, 16, 2, 12, 24, 44, 25	R, 26, 27, 41, 15, 36, 39
	XXVI	F, 29, 11b, 16, 2, 12, 24, 44, 25	R, 26, 27, 41, 15, 36, 39
	XXVIII	F, 29, 11b, 16, 2, 12, 24, 44, 25	R, 26, 27, 41, 15, 36, 39
Intra8	VIII	F, 16, 35, 12, 37	R, 33, 7, 15, 36
Intra13	XXIX	F, 11b, 16, 35, 37, 24, 25	R, 26, 27, 7b, 38, 36, 40b
	XXXII	F, 29, 11b, 16, 35, 35b, 12, 37, 24, 25	R, 26, 26b, 27, 7b, 38, 36, 40b
	XXXIII	F, 29, 11b, 16, 35, 12, 24, 25	R, 26, 27, 7b, 38, 36, 40b
Intra14	XIV, XXX	F, 29, 11b, 16, 35, 37	R, 33, 15, 36, 39
	XXXI	F, 29, 11b, 16, 35, 37, 37b	R, 33, 15, 36, 39
Intra20	XX	F, 16, 35, 12, 37, 24, 25	R, 26, 27, 7, 7b, 38, 36
Intra22	XXII	F, 29, 11, 76, 16, 12, 37	R, 30, 7b, 15, 75, 74
Intra23	XXIII	F, 29, 11, 76, 16, 12, 24, 25	R, 30, 26, 27, 7b, 15, 75, 74
Intra24	XXIV	F, 16, 35, 12, 37, 24, 25	R, 26, 27, 7, 7b, 38, 36
Intra25	XXV	29, 11, 76, 16, 12, 37, 24, 25, 83	30, 84, 26, 31, 7, 15, 75, 74
Intra27	XXVII	F, 29, 11b, 16, 35, 12, 37, 24, 44, 53, 55	R, 52, 54, 27, 7b, 38, 36, 40b
Intra3a		2	
Intra13		29	
Intra14		29	
Fung1		29, 11, 16, 12	81, 80
Fung2		F, 29, 11, 67	
Fung3		11, 16, 12	81, 79
Fung4		29, 11, 16, 12	81, 80
Fung5		29, 11, 68	30
Fung6		29, 11, 68	30
Fung7		16, 12	81, 80
Fung8		29, 11, 68	
Fung9		11, 68	
Fung10		11, 68	
Fung11		29, 16, 12	81, 15
Fung12		16, 12	81, 15

Table 4 Sequencing primers used for the mitochondrial rRNA large subunit gene (mtLSU), which were newly designed in this study. Other sequencing primers were published by Börstler *et al.* (2008) and Raab *et al*. (2005).

RNL primer	Sequence
26b	CTAGTGCAAGTAGGCCTTCT
35b	TAACCCCTTAACGACCACAC
37b	CAGCTGGGCTAAGAATGCTG
40b	CCTGCTTTGTAAGTCTACCT
67	AGATTTAACACCCGCAGCCT
68	GTCTGCCTTGTTTCAATTAC
74	CTTGCTCTTCATCTCATAGTC
75	AGATAATCTCCTTACCTTAC
76	TCTTCCTTCCGAGATAGGAG
79	GACCATCTAACATCTTACAG
80	CCTGCTGATTGTCCCTATGT
81	TGCCTTAGAGGCCGCTGCTT
83	CCCTAACAACATGTCTGTAT
84	GGCGAGAATATACTCCACCT

7. Phylogenetic analyses

Complete mtLSU sequences were analysed phylogenetically using distance, parsimony and maximum likelihood criteria as implemented in PAUP* 4.0b10 (Swofford, 2001). Neighbour joining or heuristic search algorithms were applied, respectively. Maximum likelihood models and parameters were estimated using Modeltest 3.5 (Posada, 2004).

8. Population genetics

Hierarchical F statistics was calculated using the 'hierfstat' package (Goudet, 2005) in the program package 'R' (http://www.R-project.org). P-values were obtained by 1000 permutations.

9. Statistical analyses

The sampling effort curve was calculated for each study site and all treatments of the agricultural sites in the program EstimateS 8.0.0 (Colwell, 2005), sample order was randomized in 100 replications.

The mtLSU haplotype diversity determined by RFLP was estimated for all study sites and within different treatments of the field experiments using the (i) Simpson's index of diversity, (ii) Shannon index, and (iii) Evenness index.

In order to investigate the influence of the environmental variables (field sites and different treatments of the agricultural sites) on the distribution of mtLSU haplotypes (RFLP or sequence data) of *G. intraradices*, canonical correspondence analyses (CCA) were conducted in CANOCO for Windows version 4.5 (ter Braak & Smilauer, 2004) using the haplotype presence or absence data for each root sample. Additionally, Monte-Carlo permutation tests were conducted using 499 random permutations in order to determine the statistical significance of the relation between the whole set of environmental variables and the mtLSU haplotypes. Forward selection was used for ranking the environmental variables in importance for determining the haplotype distribution in case of more than two variables.

Relationships between geographical distances and similarities of RFLP types of *G. intraradices* of the experimental plots in Frick and Tänikon were examined using the Mantel test with 99 999 permutations. The software PASSaGE (Rosenberg, 2001) was used to conduct these analyses.

IV. Results

1. Amplification success

MtLSU PCR products of *G. intraradices* (sensu GLOM A-1) were detected in 99 out of 127 root samples analysed in this study (see Table 5). PCR product lengths ranged from 1 500 to 3 293 bp for *G. intraradices* (Table 6), sizes above 4 000 bp were exclusively reached by non-target fungi (not shown). Twenty-one samples contained mixtures of up to three different haplotypes of *G. intraradices*. MtLSU PCR products from other fungi were found in 20 samples. Combinations of both products from other species and from *G. intraradices* occurred in four samples. Although all samples from Landskron and Ramosch were expected to be positive for *G. intraradices* due to positive ITS results, no mtLSU PCR product could be obtained from one sample from Ramosch.

Table 5 Number and distribution of restriction fragment length polymorphism (RFLP) patterns of *Glomus intraradices* ("Intra types") and other fungi ("Fung types") detected in the Tänikon field site (three treatments: no-tillage, chisel and tillage), the Frick field site (two treatments: chisel and tillage) and the semi-natural grasslands of Landskron and Ramosch.

Intra type	**Tänikon**				**Frick**			**Landskron**	**Ramosch**
	No-tillage	**Chisel**	**Tillage**	**∑**	**Chisel**	**Tillage**	**∑**		
Intra1	3	11	12	26					
Intra2a		2		2					
Intra3a		2	1	3	7	5	12		
Intra5a			1	1					
Intra6	1			1				2	
Intra8		2		2					
Intra13								13	16
Intra14								14	11
Intra20		1		1					
Intra22					1		1		
Intra23						1	1		
Intra24					4	1	5		
Intra25					3	8	11		
Intra27								2	
Total number of RFLP patterns of all Intra types	4	18	14	36	15	15	30	31	27
∑ different Intra types	2	5	3	7	4	4	5	4	2
Positive Intra type-root samples/root sample number (%)	4/16 (25)	13/16 (81)	12/16 (75)	29/48 (60)	11/16 (69)	13/16 (81)	24/32 (75)	21/21 (100)	25/26 (96)
Fung type									
Fung1	1			1					
Fung2	1			1					
Fung3	5	3	3	11					
Fung4	3			3					
Fung5			1	1					
Fung6			1	1					
Fung7			1	1					
Fung8						1	1		
Fung9						1	1		
Fung10						1	1		
Fung11									1
Fung12									1
Total number of RFLP patterns of all Fung types	10	3	6	19	0	3	3	0	2
∑ different Fung types	4	1	4	7	0	3	3	0	2
Positive Fung type-root samples/root sample number (%)	8/16 (50)	3/16 (19)	5/16 (31)	16/48 (33)	0/16 (0)	2/16 (13)	2/32 (6)	0/21 (0)	2/26 (8)
Root samples containing Intra - and Fung types/root sample number (%)	0/16 (0)	0/16 (0)	1/16 (6)	1/48 (2)	0/16 (0)	1/16 (6)	1/32 (3)	0/21 (0)	2/26 (8)

Table 6 Sequence structure of the mitochondrial rRNA large subunit gene (mtLSU) rDNA from root-colonizing *Glomus intraradices* obtained using primers RNL-29/RNL-30 (excluding primer annealing sites)

Origin and plant samples	Clones (accession number)	Intra type	Haplotype	Introns			Exon region	Fragment length (bp)
				Pos. 1 (-type,bp)	Pos. 2 (-type, bp)	Pos. 3 (-type, bp)	Complete/parts between intron positions (bp)	
Tänikon (Switzerland)								
15/1 (*Zea mays*)	FN377576	6	XVIII	1-9, 453	2-4, 303	3-3, 944	1067/363, 666, 38	2767
12/2 (*Zea mays*)	FN377577	8	VIII	No	2-1, 401	No	1071/363, 670, 38	1472
12/4 (*Zea mays*)	FN377578	3a	XIX	No	2-3, 389	No	1071/363, 670, 38	1460
18/2 (*Zea mays*)	FN377579	20	XX	No	2-1, 401	3-8, 802	1071/363, 670, 38	2274
18/4 (*Zea mays*)	FN377580	2a	XXI	No	2-3, 389	3-8, 802	1071/363, 670, 38	2262
24/1 (*Zea mays*)	FN377581	1	I	1-1, 1057	2-1, 401	No	1071/363, 670, 38	2529
24/2 (*Zea mays*)	FN377582	5a	V	1-3, 1094	2-3, 389	No	1071/363, 670, 38	2554
24/3 (*Zea mays*)	FN377583	1	I	1-1, 1057	2-1, 401	No	1071/363, 670, 38	2529
Frick (Switzerland)								
W7/4 (*Zea mays*)	FN377584	22	XXII	1-10, 795	No	No	1069/366, 665, 38	1864
O4/4 (*Zea mays*)	FN377585	23	XXIII	1-10, 795	No	3-1, 809	1069/366, 665, 38	2673
	FN377586	24	XXIV	No	2-6, 339	3-1, 809	1071/363, 670, 38	2219
O7/3 (*Zea mays*)	FN377587	3a	III	No	2-2, 414	No	1071/363, 670, 38	1485
W9/1 (*Zea mays*)	FN377588†	25	XXV	1-10, 795	No	3-9, 1389	1069/366, 665, 38	3253
Landskron (France)								
L5 (*Bromus erectus*)	FN377589	6	XXVI	1-11, 447	2-4, 303	3-3, 944	1067/363, 666, 38	2761
L8 (*Festuca pratensis*)	FN377590	27	XXVII	1-7, 444	2-4, 303	3-2, 1356	1067/363, 666, 38	3170
L9 (*Festuca pratensis*)	FN377591	14	XIV	1-7, 444	2-4, 303	No	1067/363, 666, 38	1814
L10 (*Festuca pratensis*)	FN377592	6	XXVIII	1-7, 444	2-4, 303	3-3, 944	1067/363, 666, 38	2758
L14 (*Medicago sativa*)	FN377593	13	XXIX	1-11, 447	2-4, 303	3-6, 739	1067/363, 666, 38	2556
	FN377594	14	XXX	1-11, 447	2-4, 303	No	1067/363, 666, 38	1817
L15 (*Medicago sativa*)	FN377595	14	XXXI	1-12, 438	2-4, 303	No	1067/363, 666, 38	1808
	FN377596	14	XXXI	1-12, 438	2-4, 303	No	1067/363, 666, 38	1808
L21 (*Inula salicina*)	FN377597	13	XXXII	1-13, 434	2-4, 303	3-6, 737	1067/363, 666, 38	2541
Ramosch (Switzerland)								
R4 (*Gentiana verna*)	FN377598	13	XXXIII	1-4, 425	2-4, 303	3-6, 737	1067/363, 666, 38	2532
R11 (*Trifolium* sp.)*S6	AM950222	14	XIV	1-7, 444	2-4, 303	No	1067/363, 666, 38	1814
	AM950223	14	XIV	1-7, 444	2-4, 303	No	1067/363, 666, 38	1814
R12 (*Trifolium* sp.)*S10	AM950224	14	XIV	1-7, 445	2-4, 303	No	1067/363, 666, 38	1815
	AM950225	14	XIV	1-7, 444	2-4, 303	No	1067/363, 666, 38	1814
R23 (Poaceae sp.)	FN377599	13	XXXII	1-13, 434	2-4, 303	3-6, 737	1067/363, 666, 38	2541
Compartment system								
C1 (*Medicago sativa*)	FN377600	13	XXIX	1-11, 447	2-4, 303	3-6, 739	1067/363, 666, 38	2556

Haplotypes and intron types were distinguished by sequence differences. Introns containing putative open reading frames for LAGLIDADG homing endonucleases are shaded. The corresponding RFLP patterns defined as 'Intra types' are listed for comparison. Inoculum for the 'compartment system' originated from the Landskron field site (Sýkorová *et al.* 2007a). For three samples, two clones of the same Intra type were sequenced. * Börstler *et al.* (2008). † PCR-product.

2. RFLP analyses

Fourteen different RFLP patterns of *G. intraradices* GLOM A-1 (referred to as Intra types in this study) could be distinguished. In addition, 12 different patterns of other fungi were found (Table 5). The sampling effort curves clearly reach saturation for both semi-natural grasslands (Fig. 6). This is not as clearly the case for the agricultural sites, although the curves begin to level off and finding an additional RFLP type would require the analysis of a considerable amount of samples. Intra type richness was highest in Tänikon followed by the

other agricultural site in Frick (Table 7, Fig. 7). The latter site also revealed higher diversity indices than the grasslands, whereas the distribution of Intra types in Tänikon was uneven, which is expressed in lower Simpson's and Shannon indices compared to the grassland in Landskron (Table 7).

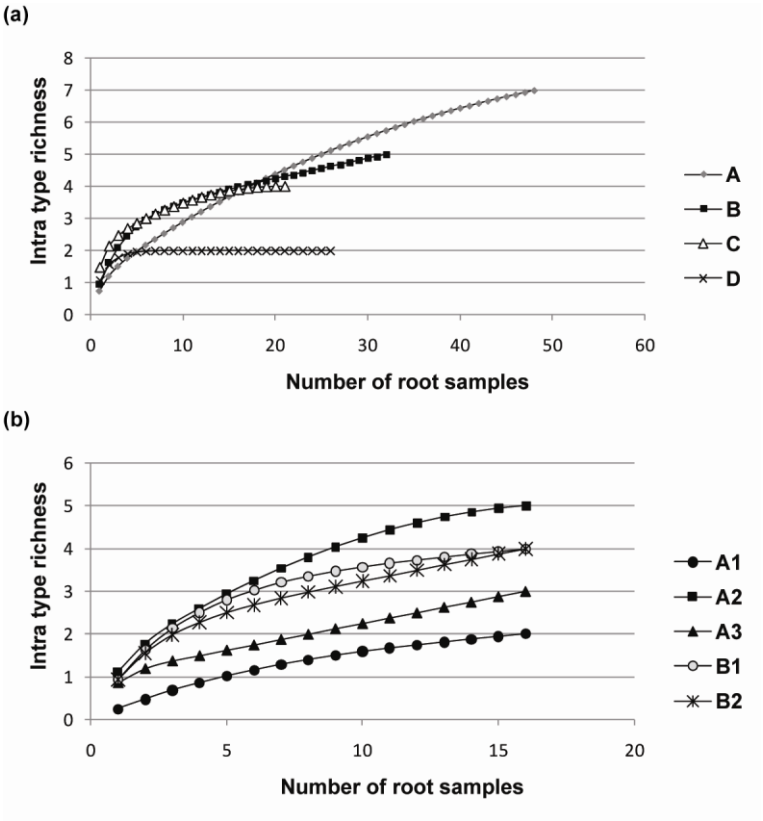


Fig. 6 (a) Sampling effort curves based on RFLP types of *Glomus intraradices* (Intra types) for the study sites Tänikon (A), Frick (B), Landskron (C) and Ramosch (D). **(b)** Sampling effort curves based on Intra types for the Tänikon treatments, no-tillage (A1), chisel (A2), tillage (A3) and the Frick treatments chisel (B1) and tillage (B2). Sample orders were randomized in 100 replications using EstimateS, version 8.0.0 (Colwell, 2005).

Table 7 Indices of diversity, richness and evenness of RFLP types (Intra types) of *Glomus intraradices* in all study sites and different treatments. F (Frick), T (Tänikon).

Fiel site or treatment	Intra type richness (S)	Simpson's index of diversity (D)	Shannon index (H)	Evenness index (E)
Tänikon	7	0.48	1.06	0.55
Frick	5	0.70	1.26	0.78
Landskron	4	0.63	1.08	0.78
Ramosch	2	0.50	0.68	0.98
T/no-tillage	2	0.50	0.56	0.81
T/chisel	5	0.62	1.19	0.74
T/tillage	3	0.27	0.51	0.46
F/chisel	4	0.71	1.21	0.87
F/tillage	4	0.64	1.06	0.77

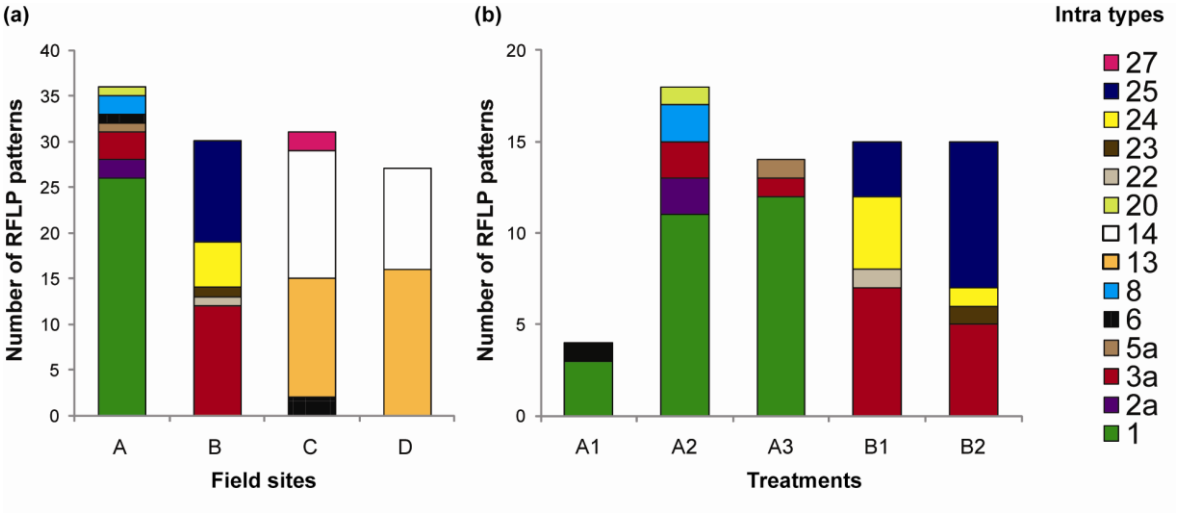


Fig. 7 (a) Total number of restriction fragment length polymorphism (RFLP) patterns of *Glomus intraradices* and different RFLP types ('Intra types') found in the Tänikon field site (A), the Frick field site (B), the Landskron grassland (C) and the Ramosch grassland (D). **(b)** Number of RFLP patterns and different Intra types found in different treatments of the Tänikon field site (A1/no-tillage, A2/chisel, A3/tillage) and the Frick field site (B1/chisel, B2/tillage).

All study sites were dominated by one or two RFLP types (see Fig. 7). RFLP patterns of these most frequently found RFLP types are shown in Fig. 8. RFLP type composition differed strongly among all study sites, except between the two grasslands: Intra13 and Intra14 were the only Intra types found frequently in Ramosch and the same types were also dominating the other semi-natural site in Landskron. Intra13 was also found in the sample from the 'compartment system', a greenhouse culture derived from inoculum from the Landskron site (Sýkorová *et al.* 2007a). Only two further Intra types were shared by more than one site: Intra3a was detected in both agricultural sites, though less frequently found in Tänikon, and Intra6 was detected in Tänikon as well as in Landskron.

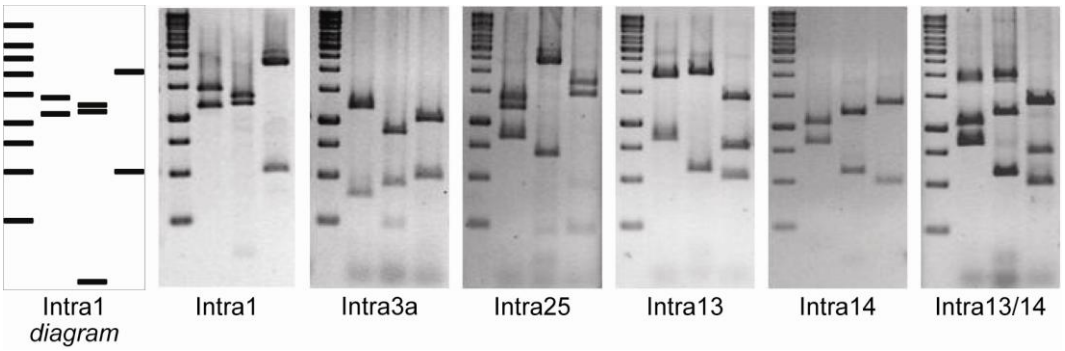


Fig. 8 Banding patterns of the most frequent *Glomus intraradices* restriction fragment length polymorphism (RFLP) patterns ("Intra types") from the Tänikon field site (Intra1), the Frick field site (Intra3a/Intra25), the grasslands Landskron and Ramosch (Intra13/Intra14). An example for a diagram made *in silico* from the sequence data is shown for Intra1. A mixed pattern is shown for Intra13/Intra14. PCR products were amplified using the primer pair RNL-29/30 for the mitochondrial rRNA large subunit gene (mtLSU) rDNA. DNA was digested using the restriction enzymes *Dra*III, *Bsa*JI, *Hin*dIII according to Börstler *et al.* (2008) and loaded in the same order per sample onto the gels. DNA ladder in left-most lane of each gel (bp): 250, 500, 750, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 5000, 6000, 8000, 10 000. Fragment standard in left-most lane of the diagram (bp): 250, 500, 750, 1000, 1500, 2000, 2500, 3000, 4000.

The highest Intra type diversity was found in the chisel treatments within each experiment in Tänikon and Frick (Table 7, Fig. 7). Interestingly, only four samples were tested positive for *G. intraradices* in the no-tillage treatment of Tänikon, which, on the contrary, revealed the highest frequency of non-*G. intraradices* RFLP type patterns (Table 5). CCA revealed that all canonical axes of Tänikon explained 11.4% of the whole variability of the data. According to the Monte-Carlo permutation test, the influence of the environmental factors was significant (P = 0.002 including and P = 0.042 excluding samples without *G. intraradices*). After ranking the factors using forward-selection, only the no-tillage treatment, characterized by many samples negative for *G. intraradices* (included as 'NoIntra' in the analyses), differed significantly from the other treatments (P = 0.002), followed by chisel (P = 0.094). If 'NoIntra' samples were excluded, none of these factors was significant (P = 0.134). The different treatments in Frick had no significant influence on the distribution of RFLP types based on the Monte-Carlo permutation test (P = 0.116). Nevertheless, the results of the CCA analyses, shown as a biplot in Fig. 9, indicate some preferences of the Intra types for the respective tillage treatment.

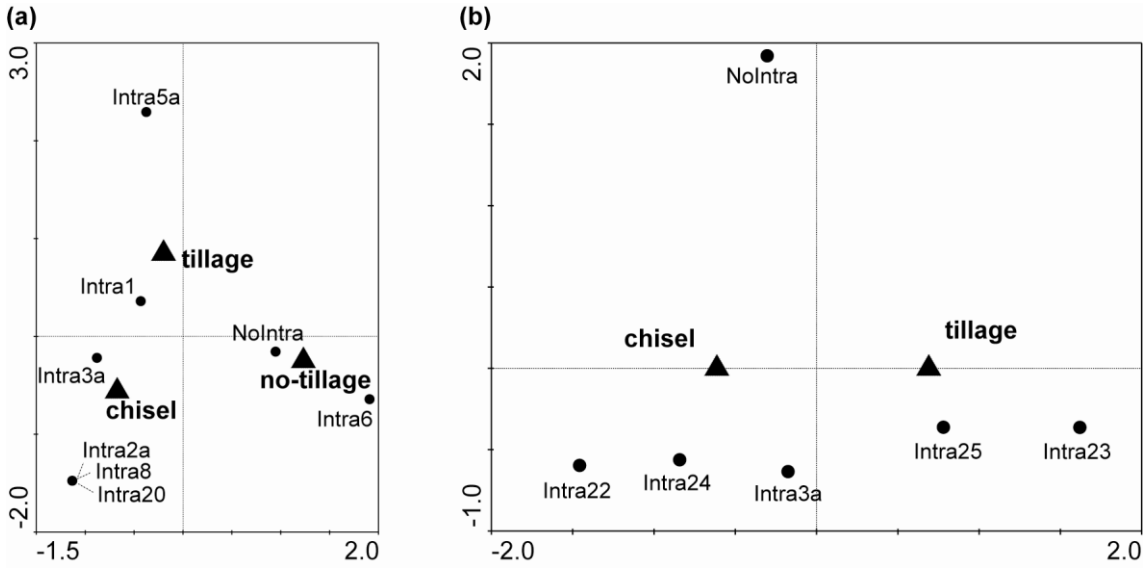


Fig. 9 CCA-biplots of RFLP type data obtained from mtLSU PCR products (using biplot scaling on symmetric distances) from **(a)** Tänikon samples and **(b)** Frick samples. The different treatments no-tillage, chisel and tillage represent the environmental variables and are shown as triangles. Filled circles indicate the position of different RFLP types of *Glomus intraradices* (Intra types) in the ordination space. Samples tested negative for *G. intraradices* were included as 'NoIntra'.

Mantel tests comparing the possible correlation of RFLP types of *G. intraradices* to spatial distance among the plots revealed no relationship between these variables for Tänikon ($r = 0.03556$; $P = 0.51566$) as well as for Frick ($r = 0.02781$; $P = 0.57692$). These results indicate that the plots located close to each other did not preferentially harbour the same *G. intraradices* RFLP types.

Hierarchical F statistics revealed significant genetic differentiation between the two arable sites which was even slightly stronger when phylogenetic outliers (see Fig. 10; sister group containing haplotypes XVII, XXII, XXIII, XXV) were excluded ($F_{ST} = 0.4631$, $P = 0.001$). Treatments within these sites were not significantly different ($F_{(site/treatment)} = 0.0627$, $P = 0.142$). In contrast, the two grasslands did not show significant differences in their RFLP types ($F_{ST} = -0.00361$, $P = 0.35$), neither were host preferences detectable ($F_{(site / host)} = 0.04519$, $P = 0.124$).

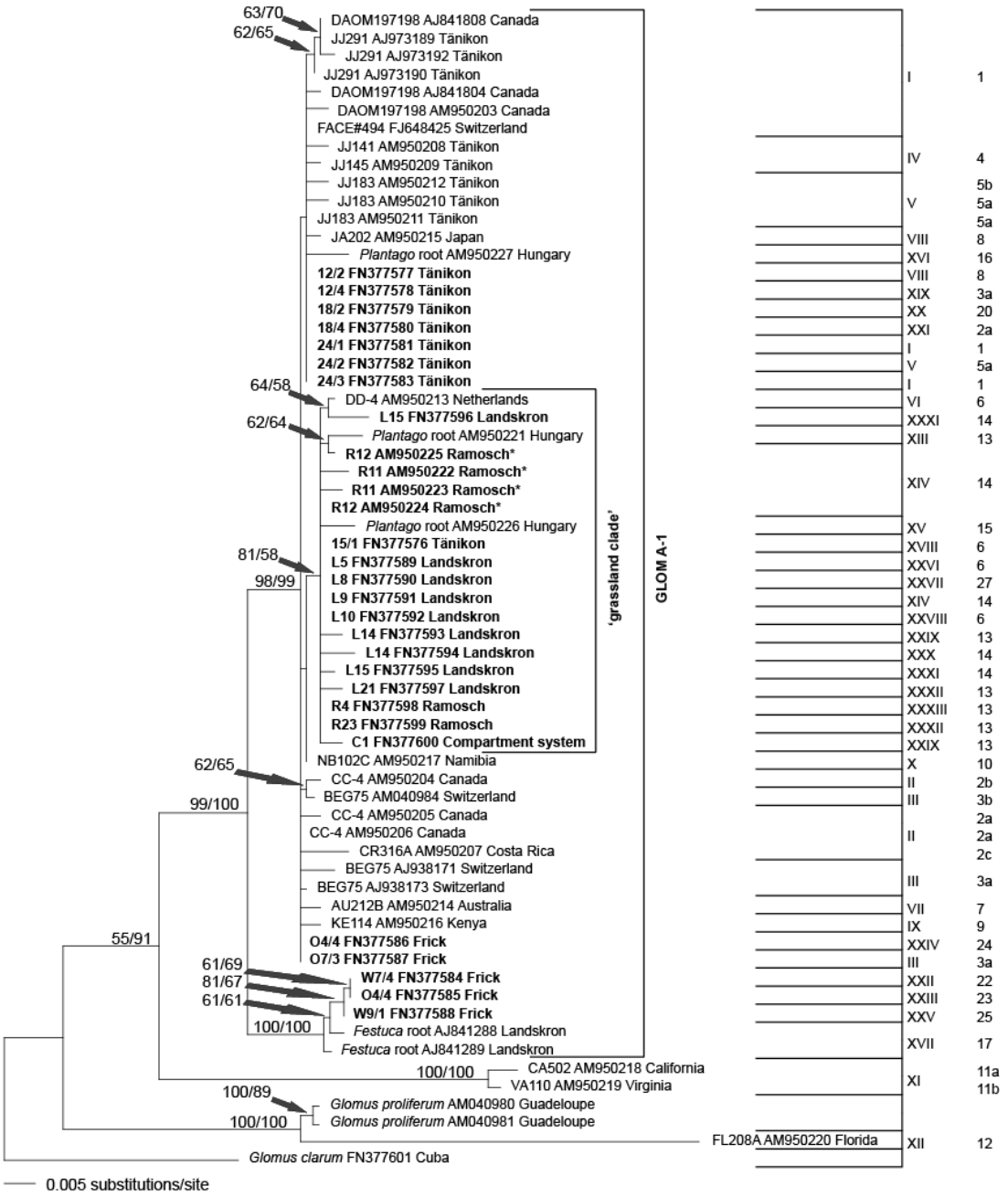


Fig. 10 Phylogenetic tree of *Glomus intraradices*, *Glomus proliferum* and *Glomus clarum*, which was used as outgroup. The phylogeny was obtained using 931 characters of mitochondrial rRNA large subunit gene (mtLSU) exon sequences by a heuristic search under the maximum likelihood criterion. Maximum parsimony bootstrap values from 1000 replicates and Neighbour-joining bootstrap values from 1000 replicates are shown at the nodes. Sample code, accession number and origin are shown in boldface for sequences from root samples originating from the arable field sites and grasslands of the present study. *G. intraradices* isolates are labelled with their identification code. Roman numerals indicate mtLSU haplotypes, Arabic numerals indicate RFLP types ('Intra types'). *Sequences from Ramosch previously analysed by Börstler *et al.* (2008).

Because of a different sampling structure and the possibility that grassland phylotypes might represent different species (see below), therefore precluding an overall population analysis, grassland and arable site samples were not analysed together. On the contrary, the fact that only the rarely found RFLP type Intra6 was shared between arable and semi-natural sites is obvious evidence supporting genetic differentiation (Fig. 7).

3. Sequence-based analyses

Although RFLP types were shared by some sites, detailed sequence analyses revealed that there was no overlap of mtLSU haplotypes among any of the study sites except between the two grasslands. All PCR products of Intra6, found in Tänikon and Landskron, were sequenced completely and differed within and between the sites in the intron at position 1 (Table 6). Intra3a which was shared by the two agricultural sites contained only an intron at position 2 (Table 6), which was sequenced for all PCR products as the only sequence region where polymorphism would be expected. Sequencing revealed that the mtLSU sequence type detected within Intra3a was the same within each field site, but they differed between the sites (Fig. 11).

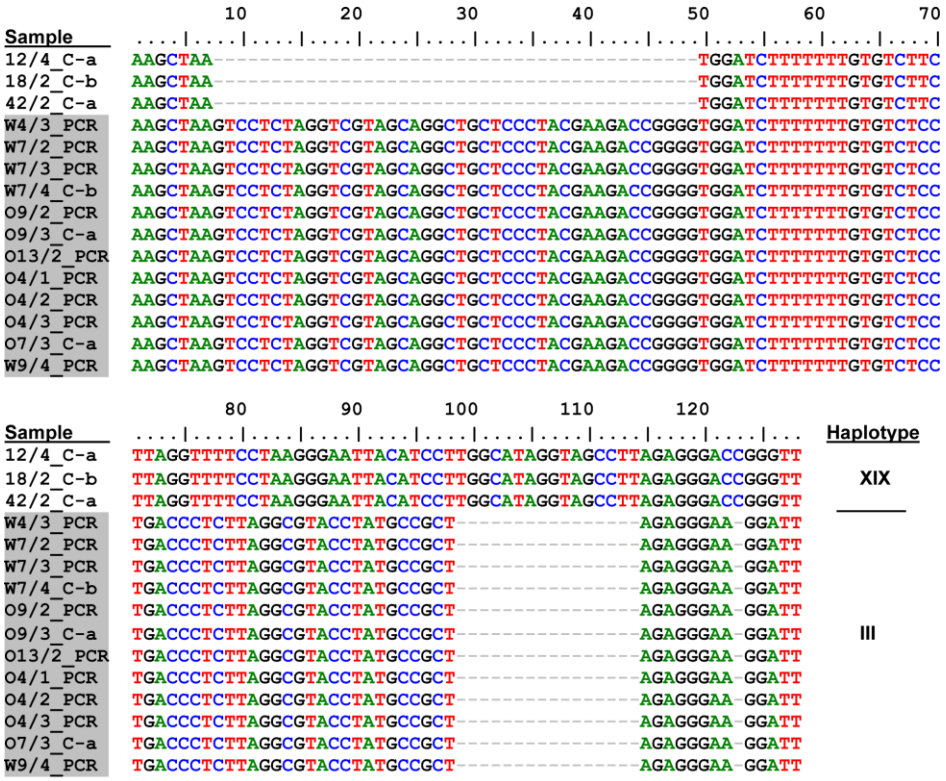
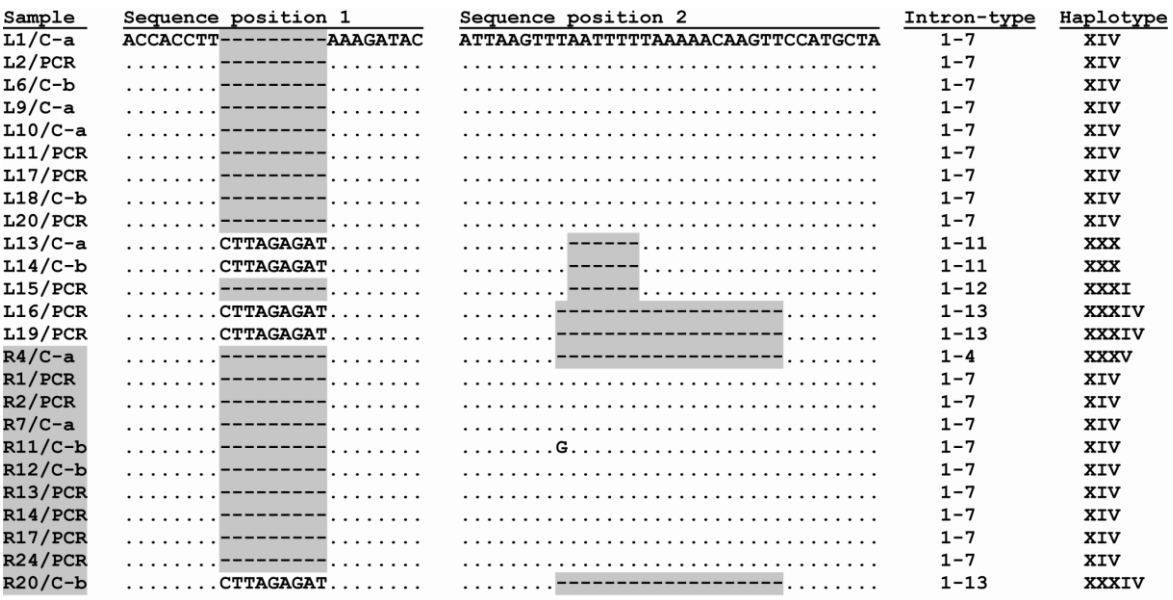


Fig. 11 Alignment of the diagnostic sequence region in the intron at position 2 of mtLSU haplotypes represented by the same RFLP type "Intra3a" from Tänikon and Frick (samples from Frick are shaded). Gaps are shown as hyphens. An unequivocal alignment of the region between positions 71 until 114 is not possible. Sample identification codes are provided in addition to the origin of sequences from clones (C) or PCR products (PCR).

Completely sequenced PCR products of Intra13 and Intra14 from five root samples each confirmed that the only variable sequence regions in these PCR products were located in the intron at position 1 (Table 6), which was expected from previous analyses of an isolate from the Netherlands and a root sample from Hungary (Börstler et al., 2008). Therefore, this gene region was sequenced from all PCR products assigned to these RFLP types. Based on these partial sequences, the two Intra types could be further separated into nine different sequence haplotypes (Figs 12 and 13). The differences consisted mainly of short indels and not of more substantial sequence differences as they were typical for different introns of the highly diverse agricultural field sites (Table 6).

Sample	Sequence position 1	Sequence position 2	Intron-type	Haplotype
L1/C-b	ACCACCTT---------AAAGATAC	ATTAAGTTTCATTTTTAAAAACAAGTTCCATGCTA	1-7	XIII
L6/C-a	........---------........	.........A..........................	1-7	XIII
L8/C-b	........---------........	.........A..........................	1-7	XIII
L9/C-b	........---------........	.........A..........................	1-7	XIII
L18/C-a	........---------........	.........A..........................	1-7	XIII
L3/C-a	........CTTAGAGAT........	.........------.....................	1-11	XXIX
L4/PCR	........CTTAGAGAT........	.........------.....................	1-11	XXIX
L5/C-a	........CTTAGAGAT........	.........------.....................	1-11	XXIX
L7/PCR	........CTTAGAGAT........	.........------.....................	1-11	XXIX
L12/PCR	........CTTAGAGAT........	.........------.....................	1-11	XXIX
L13/C-b	........CTTAGAGAT........	.........------.....................	1-11	XXIX
L14/C-a	........CTTAGAGAT........	.........------.....................	1-11	XXIX
L21/C-a	........CTTAGAGAT........	........------------------........	1-13	XXXII
R3/PCR	........CTTAGAGAT........	........------------------........	1-13	XXXII
R6/PCR	........CTTAGAGAT........	........------------------........	1-13	XXXII
R8/PCR	........CTTAGAGAT........	........------------------........	1-13	XXXII
R9/PCR	........CTTAGAGAT........	........------------------........	1-13	XXXII
R10/PCR	........CTTAGAGAT........	........------------------........	1-13	XXXII
R15/PCR	........CTTAGAGAT........	........------------------........	1-13	XXXII
R16/PCR	........CTTAGAGAT........	........------------------........	1-13	XXXII
R18/PCR	........CTTAGAGAT........	........------------------........	1-13	XXXII
R19/C-a	........CTTAGAGAT........	........------------------........	1-13	XXXII
R20/C-a	........CTTAGAGAT........	........------------------........	1-13	XXXII
R21/PCR	........CTTAGAGAT........	........------------------........	1-13	XXXII
R22/PCR	........CTTAGAGAT........	........------------------........	1-13	XXXII
R23/C-a	........CTTAGAGAT........	........------------------........	1-13	XXXII
R25/C-a	........CTTAGAGAT........	........------------------........	1-13	XXXII
R26/C-a	........CTTAGAGAT........	........------------------........	1-13	XXXII
R4/C-b	........---------........	........------------------........	1-4	XXXIII
C1/C-a	........CTTAGAGAT........	.........------...........C........	1-11	XXIX
AM950221	......C.---------........	....	1-7	XIII

Fig. 12 Alignment of two diagnostic sequence regions in the intron at position 1 of mtLSU PCR products represented by the same RFLP type "Intra13" from Landskron and Ramosch. Conservation is viewed by plotting identities to a standard as dot, gaps are given as shaded hyphens. Samples from the Ramosch grassland are shaded. Two additional samples from the compartment system based on inoculum from Landskron (C1) and from Hungary (AM950221; Börstler *et al.* 2008) are given for comparison. Sample identification codes are given in addition to the origin of sequences from clones (C) or PCR products (PCR). Three changes of a single base occurred for samples amplified with *Taq* polymerase (C1, AM950221), only one change of a single base occurred for sample L1 amplified with Phusion polymerase. All these single changes are most probably caused by polymerase or sequencing errors.



Sample	Sequence position 1	Sequence position 2	Intron-type	Haplotype
L1/C-a	ACCACCTT---------AAAGATAC	ATTAAGTTTAATTTTTAAAAACAAGTTCCATGCTA	1-7	XIV
L2/PCR	........---------........	.....	1-7	XIV
L6/C-b	........---------........	.....	1-7	XIV
L9/C-a	........---------........	.....	1-7	XIV
L10/C-a	........---------........	.....	1-7	XIV
L11/PCR	........---------........	.....	1-7	XIV
L17/PCR	........---------........	.....	1-7	XIV
L18/C-b	........---------........	.....	1-7	XIV
L20/PCR	........---------........	.....	1-7	XIV
L13/C-a	........CTTAGAGAT........	.........------.....	1-11	XXX
L14/C-b	........CTTAGAGAT........	.........------.....	1-11	XXX
L15/PCR	........---------........	.........------.....	1-12	XXXI
L16/PCR	........CTTAGAGAT........	.........----------....	1-13	XXXIV
L19/PCR	........CTTAGAGAT........	.........----------....	1-13	XXXIV
R4/C-a	........---------........	.........----------....	1-4	XXXV
R1/PCR	........---------........	.....	1-7	XIV
R2/PCR	........---------........	.....	1-7	XIV
R7/C-a	........---------........	.....	1-7	XIV
R11/C-b	........---------........	........G.....	1-7	XIV
R12/C-b	........---------........	.....	1-7	XIV
R13/PCR	........---------........	.....	1-7	XIV
R14/PCR	........---------........	.....	1-7	XIV
R17/PCR	........---------........	.....	1-7	XIV
R24/PCR	........---------........	.....	1-7	XIV
R20/C-b	........CTTAGAGAT........	.........----------....	1-13	XXXIV

Fig. 13 Alignment of two diagnostic sequence regions in the intron at position 1 of mtLSU PCR products represented by the same RFLP type "Intra14" from Landskron and Ramosch. Conservation is viewed by plotting identities to a standard as dot, gaps are given as shaded hyphens. Samples from the Ramosch grassland are shaded. Sample identification codes are given in addition to the origin of sequences from clones (C) or PCR products (PCR). Only one change of a single base occurred for sample R11 previously amplified with *Taq* polymerase by Börstler *et al.* (2008) and is most probably due to a polymerase or sequencing error.

In contrast to the arable sites, the grassland samples were thus exhaustively analysed by sequencing (see also below for the remaining RFLP type Intra27), and the sequence types could therefore be included into statistical analyses. The distribution of all 12 different mtLSU haplotypes based on sequence data across the grasslands is shown in Fig. 14: haplotype XIV was found frequently in both sites. Two further haplotypes were shared: haplotype XXXII was most frequently found in Ramosch, but only rarely detected in Landskron and haplotype XXXIV was found rarely in both grasslands. The CCA revealed that the environmental variables explained 8.5% of the total variance and according to the Monte-Carlo permutation test the influence of the field sites on haplotype distribution was significant (P = 0.002). The results of the CCA analysis are shown as a biplot in Fig. 15.

Contrary to the RFLP analysis, the two grasslands sites were significantly differentiated based on sequence haplotype data (F_{ST} = 0.17448, P = 0.01). Host preference, however, again only had a weak influence on population structure ($F_{(site\,/\,host)}$ = 0.04519, P = 0.124).

Generally, at least one example of each different RFLP type of *G. intraradices* in every study site was sequenced completely for the determination of the respective mtLSU haplotype based on sequence data (Table 6).

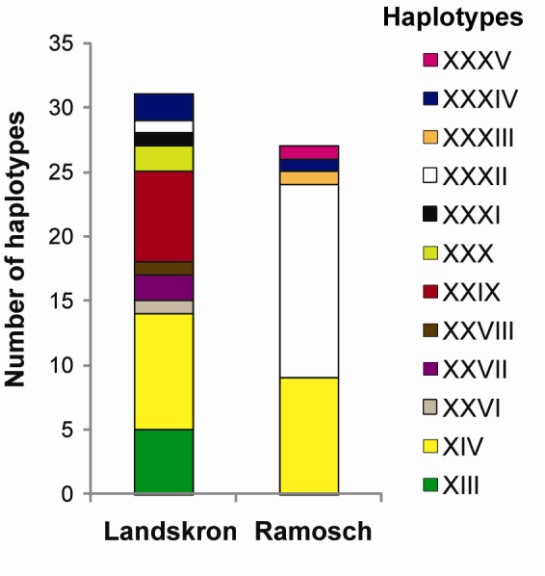


Fig. 14 Number and distribution of different mtLSU haplotypes of *Glomus intraradices* found in the grasslands in Landskron and Ramosch.

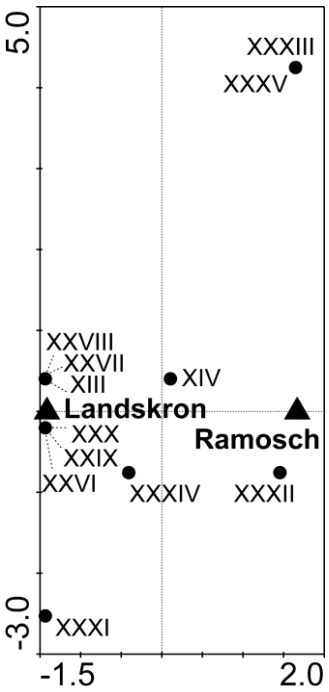


Fig. 15 CCA-biplot of mtLSU haplotype data (using Hill's scaling focused on inter-species distances) from Landskron and Ramosch. The field sites represent the environmental variables and are shown as triangles. Filled circles indicate the position of 12 different haplotypes of *Glomus intraradices* in the ordination space, which are given in Roman numerals. Only one sample from Ramosch was tested negative for *G. intraradices* and was excluded from the analysis.

4. Phylogenetic analyses

All completely sequenced mtLSU PCR products were compared phylogenetically with sequences from isolates of *G. intraradices* and previously analysed root samples generated by Raab *et al.* (2005) and Börstler *et al.* (2008). A phylogenetic tree was inferred (Fig. 10). All haplotypes from the two grasslands and the 'compartment system' derived from the Landskron site were falling into one subclade, which also comprises two haplotypes from two root samples from Hungary and only one haplotype of an isolate from a Dutch dry dune grassland. Only one haplotype from the agricultural sites was found in this group. Interestingly, this unique haplotype was found in the no-tillage treatment of the Tänikon site. Because of a small number of unique sequence positions, this 'grassland subclade' was consistently recovered throughout all phylogenetic analyses although it only received moderate bootstrap support, and was not recovered as sister group to the remaining isolates including DAOM197198 (which is subject of the genome sequencing project of *G. intraradices*) and FACE494, of which the complete mitochondrial genome was sequenced by Lee & Young (2009).

A clearly distinct sister clade to these two groups comprised sequences found in the Landskron grassland by Raab *et al.* (2005), using different primers, and the Frick arable site. It is currently impossible to decide whether this lineage or even the grassland clade constitute different biological species, but to avoid confounding analyses of intraspecific genetic structure variation we excluded these groups or analysed them separately in some analyses to assess whether this would affect our main conclusions.

5. The specific approach to detect haplotype I

A nested PCR approach designed to specifically detect mtLSU haplotype I uses characteristic sequence regions of the introns at position 1 and 2, and results in a 320-bp DNA fragment of the first intron, which is easy to amplify (see Fig. 5). The PCR product can be sequenced directly without any cloning steps. This approach was applied on the DNA extracts from all study sites. Positive PCR products were only obtained in Tänikon samples also containing the RFLP type Intra1, which corresponds to haplotype I. In other words, the specific approach for haplotype I confirmed the results of the general PCR-RFLP approach for *G. intraradices* in each case and vice versa. All PCR products were additionally confirmed by sequencing. So far, only the haplotype X of the isolate NB102C (Börstler et al., 2008)

shows the same sequence for the final PCR product, but an amplification of this haplotype would not be possible because of a missing primer annealing site.

V. Discussion

By resolving intraspecies mtLSU haplotypes of *G. intraradices*, this study is the first one to address the genetic structure of AMF populations directly in the symbiotic state, that is, without the need to rely on spores or fungal isolates obtained from the field. The method we used here was shown to work reliably: the two different approaches (i) using group-specific primers and RFLP, and (ii) using haplotype-specific primers, respectively, provided concordant results. The specificity of the group-specific primers was high enough to assure that 84% of the amplified sequences were expected targets. Although some mtLSU RFLP types or sequence haplotypes might be further resolved by multilocus markers (Croll et al., 2008b), our RFLP data demonstrate considerable genetic structure among field sites, which was refined by sequencing of diagnostic regions in a subset of samples. In contrast to our previous studies (Raab *et al.*, 2005; Börstler *et al.*, 2008), we used mainly a proofreading DNA polymerase which considerably reduced sequencing noise in our datasets. In samples, containing several RFLP types, minority types were visible as faint bands, demonstrating that the approach is suitable for diagnosing multiple types without the need for cloning / sequencing approaches which may bear the high possibility of overseeing minority types.

When assigning mtLSU haplotypes to the strains used by Koch et al. (2004) and Croll et al. (2008b), it becomes clear that many of the significant effects of AMF genotypes on symbiotic properties and growth parameters in these studies are due to different mtLSU haplotypes, indicating that these haplotypes are functionally relevant categories. On the contrary, mtLSU types do not necessarily represent separate biological species, because Croll et al. (2009) demonstrated that representative strains of all three mtLSU haplotypes found by these authors in the Tänikon field site formed anastomoses with each other and exchanged genetic markers.

We analysed samples from the same field site in Tänikon that was assessed by Croll et al. (2008b) with a different approach and the comparison of the datasets offers interesting insights. In contrast to these authors, who characterized isolates of ROCs originally obtained in trap cultures from this site in 1999 (Jansa et al., 2002a), we directly used the colonized roots from the same plots, although 8 years later. Croll et al. (2008b) found three mitochondrial haplotypes, whereas we recovered two of those (I, V) and five additional ones

(VIII, XVIII, XIX, XX, XXI). Haplotype VIII was identical to an isolate JA202 originating from Japan (Börstler et al., 2008), further supporting a wide geographical distribution of certain haplotypes.

Interestingly, haplotype I, which already dominated among the isolates from this site in 1999, was also the most frequently found haplotype within the roots in 2007. The genome sequencing isolate from Canada (DAOM197198) also belongs to this haplotype, as well as isolates from other sites (Eschikon and Changins) in Switzerland (Croll et al., 2008b). Overall, this would indicate that haplotype I is particularly well adapted to both arable soils and root organ cultures and geographically widespread. Therefore, it came as a surprise that this haplotype was not at all present in the second arable field site in Frick we analysed. These results were confirmed with a primer specifically designed for haplotype I, ruling out any possibility that the haplotype was merely concealed by other, more abundant ones, which may happen in approaches using general primers.

Our findings are in contrast to the studies by Koch et al. (2004) and Croll et al. (2008b), which suggested that the global genetic diversity of *G. intraradices* is not higher than the diversity within the Swiss field site these authors analysed. As already shown by Börstler et al. (2008) based on isolates in open pot culture, the mtLSU haplotype diversity worldwide is considerably higher. Now we also demonstrate that direct detection from roots allows to detect a higher haplotype richness than cultivation procedures, which involve several bottlenecks and biases such as the initial propagation in pot cultures using different host plants (‘trap culturing’) and establishment of ROCs from single spores. As our samples and those of Croll et al. (2008b) were not taken in the same year, we cannot completely exclude that temporal fluctuations among the haplotypes occurred, but there is no reason to assume that the RFLP type richness in Tänikon has more than doubled only due to the different sampling time. The sampling effort curve (Fig. 6) suggests that the richness may in fact even be higher.

Although we could not find statistically-significant differences among the tillage treatments of the arable sites, obvious tendencies of the treatments to alter the community structure of *G. intraradices* were observed. Similarly, Rosendahl & Matzen (2008) reported that agricultural practice affects both the abundance and the population structure of different AMF species. The strongly reduced frequency of *G. intraradices* in the no-tillage treatment of Tänikon, also expressed by a lower haplotype richness, might be explained by a replacement of *G. intraradices* by other species better adapted to these environmental condition. On the

AMF species level Jansa et al. (2003) already showed that tillage affected the community structure in the field experiment of Tänikon.

An unexpected finding of our study was the genetic differentiation between the two arable field sites, which were situated at a distance of 67 km, indicating that soil chemistry and geographic factors may have a stronger influence on population structuring in agriculturally used soils than previously recognized. It will be intriguing to specifically address these possible relations in future studies. Nevertheless, the data suggest that even in arable soils, only a subset of the genotypes of *G. intraradices* may be easily recovered by culturing.

In contrast to the arable sites, the grasslands contained a completely different set of mtLSU haplotypes. Only one haplotype XIII from the semi-natural sites was previously detected in a root sample from a mine spoil in Hungary (Börstler et al., 2008). In fact, the 'grassland types' could consistently be phylogenetically distinguished from the arable soil types, although only by a few different base pairs, which was impossible by nuclear-encoded rDNA (Sýkorová *et al.* 2007a,b). This may either mean that the 'grassland types' constitute a genetically different ecotype, or that they even are a previously unrecognized ('cryptic') species. Interestingly, the only 'grassland type' detected in an arable site (haplotype XVIII) was found in the no-tillage treatment, which is the least disturbed treatment of the two field trials. In contrast to some haplotypes from arable fields, which were reported to form hyphal cross-links (anastomoses; Croll et al., 2009), and therefore can be safely assumed to belong to one biological species, the grassland types have not been analysed yet. This is complicated by the fact that they are apparently difficult to cultivate, even in conventional open pot culture, and therefore are strongly under-represented among cultivated isolates. In any case, the strong differentiation between arable and semi-natural sites is a finding that will have considerable impact, for example, on conservation biology, because it renders it unlikely that semi-natural grasslands act as diversity reservoirs for arable sites for *G. intraradices*. Detailed studies of genotypes in arable soils and neighbouring grasslands, including intensively used farmland meadows, will be necessary to answer this question conclusively. However, the occurrence of a 'grassland type' in a greenhouse culture obtained from a grassland site ('compartment system'), demonstrates that at least some of these haplotypes are in principle capable of colonizing roots under high-disturbance conditions. The detection of members of this clade in degraded but non-arable ecosystems from Hungary supports this view. Further analyses of other habitats will elucidate under which range of environmental conditions these fungi thrive.

The example of the sister group of haplotypes outside the clade containing arable land and grassland haplotypes in Fig. 10, which we consider neither to belong to the arable land or grassland groups, shows that even similar haplotypes may occur in very different environments. The 'grassland types' are relatively similar to each other in their mtLSU intron sequences and are mainly distinguished only by short indels not easily picked up by RFLP. Therefore, the two grassland sites were not differentiated based on mtLSU RFLP data. Using sequences of diagnostic regions, however, we detected significant differentiation. Still, it seems remarkable that the two grasslands were the only sites in our study to share sequence haplotypes, considering that they are 230 km apart on different sides of the Alps and strongly differing with regard to edaphic conditions and plant communities.

Although it was shown that species-level richness is not necessarily reduced to one or two species in all arable soils (Hijri et al., 2006), it is generally expected to be lower than in semi-natural soils (Helgason et al., 1998a; Oehl et al., 2003). In contrast, the data presented here show that *G. intraradices* mtLSU RFLP types are in fact more diverse in arable sites and allow to hypothesize that disturbance promotes genotypic diversity in some adapted AMF species, rather than reducing it. Higher intraspecific diversity could also be correlated with higher abundance of the species in a given setting. Testing this would in addition require assessing species-level diversity qualitatively and quantitatively. In this context it is interesting that Munkvold et al. (2004) proposed that even in systems of low species diversity, considerable functional diversity is still maintained.

Studies of populations of ectomycorrhizal fungi (Kretzer et al., 2005; Geml et al., 2006; Carriconde et al., 2008) have frequently revealed unexpectedly high diversity and 'cryptic' species, but few generalizations seem to be possible about their biogeography, which seems to be related to specific ecological strategies of the respective species (Bruns & Kennedy, 2009). On the contrary, studying populations of phytopathogenic fungi has provided a better understanding of their epidemiology and the evolutionary and ecological processes involved (e.g. (Hovmøller et al., 2008)).

By providing the first direct insights into the genetic structure of field populations of the Glomeromycota, an ecologically important group of mutualistic symbionts, our study opens up a new range of questions that can be addressed now. It demonstrates that *G. intraradices* GLOM A-1, possibly one of the most frequently occurring fungal species or species complexes in the world, is strongly structured genetically with regard to habitats and environmental conditions. It is a first step towards a better understanding of the processes generating and maintaining diversity in AMF and their reciprocal interaction with plants.

VI. Acknowledgements

This work was funded by grants by the Swiss National Science Foundation to the senior author (grant numbers 3100A0-109466 and 31003A-124966), which is gratefully acknowledged. Z.S. was supported by a grant of the Ministry of Education, Youth and Sports of the Czech Republic, no. 1M0571 Research Centre for Bioindication and Revitalization. The authors acknowledge Thomas Boller and Andres Wiemken at the Botanical Institute for continuing support, Jan Jansa for providing access to the root samples from Tänikon, Paul Mäder for providing access to the samples from Frick, Maike Krauß for sampling, Kurt Ineichen for cultivating *Glomus clarum*, Sibel Gürler for performing DNA extractions, Monika Messmer for helpful discussions, and the technical staff at Hebelstrasse for making things work.

Bibliography

Agapow, P.M., Bininda-Emonds, O.R.P., Crandall, K.A., Gittleman, J.L., Mace, G.M., Marshall, J.C., Purvis, A., 2004. The impact of species concept on biodiversity studies. Q. Rev. Biol. 79, 161-179.

Akiyama, K., Hayashi, H., 2006. Strigolactones: Chemical signals for fungal symbionts and parasitic weeds in plant roots. Ann. Bot. London 97, 925-931.

Anderson, I.C., Parkin, P.I., 2007. Detection of active soil fungi by RT-PCR amplification of precursor rRNA molecules. J. Microbiol. Meth. 68, 248-253.

Antoniolli, Z.I., Schachtman, D.P., Ophel-Keller, K., Smith, S.E., 2000. Variation in rDNA ITS sequences in *Glomus mosseae* and *Gigaspora margarita* spores from a permanent pasture. Mycol. Res. 104, 708-715.

Appoloni, S. 2006. *Molecular community analysis of arbuscular mycorrhizal fungi in roots of geothermal soils in Yellowstone National Park (USA)*. MSc thesis. University of Basel, Basel.

Appoloni, S., Lekberg, Y., Tercek, M.T., Zabinski, C.A., Redecker, D., 2008. Molecular community analysis of arbuscular mycorrhizal fungi in roots of geothermal soils in Yellowstone National Park (USA). Microb. Ecol. 56, 649-659.

Avio, L., Cristani, C., Strani, P., Giovannetti, M., 2009. Genetic and phenotypic diversity of geographically different isolates of *Glomus mosseae*. Can. J. Microbiol. 55, 242-253.

Bécard, G., Fortin, J.A., 1988. Early events of vesicular arbuscular mycorrhiza formation on Ri T-DNA transformed roots. New Phytol. 108, 211-218.

Bécard, G., Pfeffer, P.E., 1993. Status of nuclear division in arbuscular mycorrhizal fungi during *in vitro* development. Protoplasma 174, 62-68.

Begerow, D., Nilsson, H., Unterseher, M., Maier, W., 2010. Current state and perspectives of fungal DNA barcoding and rapid identification procedures. Appl. Microbiol. Biot. 87, 99-108.

Belfort, M., Roberts, R.J., 1997. Homing endonucleases: keeping the house in order. Nucleic Acids Res. 25, 3379-3388.

Bentivenga, S.P., Morton, J.B., 1996. Congruence of fatty acid methyl ester profiles and morphological characters of arbuscular mycorrhizal fungi in Gigasporaceae. Proc. Natl. Acad. Sci. USA 93, 5659-5662.

Berner, A., Hildermann, I., Fliessbach, A., Pfiffner, L., Niggli, U., Mader, P., 2008. Crop yield and soil fertility response to reduced tillage under organic management. Soil Till. Res. 101, 89-96.

Besserer, A., Puech-Pages, V., Kiefer, P., Gomez-Roldan, V., Jauneau, A., Roy, S., Portais, J.C., Roux, C., Bécard, G., Sejalon-Delmas, N., 2006. Strigolactones stimulate arbuscular mycorrhizal fungi by activating mitochondria. Plos Biol. 4, 1239-1247.

Bever, J.D., Wang, M., 2005. Arbuscular mycorrhizal fungi - Hyphal fusion and multigenomic structure. Nature 433, E3-E4.

Bhattacharya, D., Reeb, V., Simon, D.M., Lutzoni, F., 2005. Phylogenetic analyses suggest reverse splicing spread of group I introns in fungal ribosomal DNA. BMC Evol. Biol. 5, 11.

Bidartondo, M.I., Redecker, D., Hijri, I., Wiemken, A., Bruns, T.D., Dominguez, L., Sersic, A., Leake, J.R., Read, D.J., 2002. Epiparasitic plants specialized on arbuscular mycorrhizal fungi. Nature 419, 389-392.

Birky, C.W., Maruyama, T., Fuerst, P., 1983. An approach to population and evolutionary genetic theory for genes in mitochondria and chloroplasts, and some results. Genetics 103, 513-527.

Birky, C.W., 1995. Uniparental inheritance of mitochondrial and chloroplast genes - Mechanisms and evolution. Proc. Natl. Acad. Sci. USA 92, 11331-11338.

Borriello, R. 2010. *"Arbuscular Mycorrhizal Fungal Diversity In Agricultural And Natural Ecosystem Soils". Looking for new molecular tools and targets for AM fungal characterization.* Università degli Studi di Torino, Torino

Börstler, B., Renker, C., Kahmen, A., Buscot, F., 2006. Species composition of arbuscular mycorrhizal fungi in two mountain meadows with differing management types and levels of plant biodiversity. Biol. Fert. Soils 42, 286-298.

Börstler, B., Raab, P.A., Thiéry, O., Morton, J.B., Redecker, D., 2008. Genetic diversity of the arbuscular mycorrhizal fungus *Glomus intraradices* as determined by mitochondrial large subunit rRNA gene sequences is considerably higher than previously expected. New Phytol. 180, 452-465.

Börstler, B., Thiéry, O., Sýkorová, Z., Berner, A., Redecker, D., 2010. Diversity of mitochondrial large subunit rDNA haplotypes of *Glomus intraradices* in two agricultural field experiments and two semi-natural grasslands. Mol. Ecol. 19, 1497-1511.

Bruns, T.D., Fogel, R., White, T.J., Palmer, J.D., 1989. Accelerated evolution of a false-truffle from a mushroom ancestor. Nature 339, 140-142.

Bruns, T.D., Szaro, T.M., Gardes, M., Cullings, K.W., Pan, J.J., Taylor, D.L., Horton, T.R., Kretzer, A., Garbelotto, M., Li, Y., 1998. A sequence database for the identification of ectomycorrhizal basidiomycetes by phylogenetic analysis. Mol. Ecol. 7, 257-272.

Bruns, T.D., Kennedy, P.G., 2009. Individuals, populations, communities and function: the growing field of ectomycorrhizal ecology. New Phytol. 182, 12-14.

Buckler, E.S.I., Ippolito, A., Holtsford, T.P., 1997. The evolution of ribosomal DNA: Divergent paralogues and phylogenetic implications. Genetics 145, 821-832.

Bunn, R., Lekberg, Y., Zabinski, C., 2009. Arbuscular mycorrhizal fungi ameliorate temperature stress in thermophilic plants. Ecology 90, 1378-1388.

Bunn, R.A., Zabinski, C.A., 2003. Arbuscular mycorrhizae in thermal-influenced soils in Yellowstone National Park. West. N. Am. Naturalist 63, 409-415.

Burleigh, S.H., Cavagnaro, T., Jakobsen, I., 2002. Functional diversity of arbuscular mycorrhizas extends to the expression of plant genes involved in P nutrition. J. Exp. Bot. 53, 1593-1601.

Burt, A., Koufopanou, V., 2004. Homing endonuclease genes: the rise and fall and rise again of a selfish element. Curr. Opin. Genet. Dev. 14, 609-615.

Calonje, M., Martin-Bravo, S., Dobes, C., Gong, W., Jordon-Thaden, I., Kiefer, C., Kiefer, M., Paule, J., Schmickl, R., Koch, M.A., 2009. Non-coding nuclear DNA markers in phylogenetic reconstruction. Plant Syst. Evol. 282, 257-280.

Carriconde, F., Gardes, M., Jargeat, P., Heilmann-Clausen, J., Mouhamadou, B., Gryta, H., 2008. Population evidence of cryptic species and geographical structure in the cosmopolitan ectomycorrhizal fungus, *Tricholoma scalpturatum*. Microb. Ecol. 56, 513-524.

Cech, T.R., 1985. Self-Splicing RNA. Biochemistry 24, 3370-3371.

Cech, T.R., 1990. Self-Splicing of Group-I Introns. Annu. Rev. Biochem. 59, 543-568.

Chevalier, B.S., Stoddard, B.L., 2001. Homing endonucleases: structural and functional insight into the catalysts of intron/intein mobility. Nucleic Acids Res. 29, 3757-3774.

Cho, Y., Qiu, Y.L., Kuhlman, P., Palmer, J.D., 1998. Explosive invasion of plant mitochondria by a group I intron. Proc. Natl. Acad. Sci. USA 95, 14244-14249.

Cho, Y.R., Palmer, J.D., 1999. Multiple acquisitions via horizontal transfer of a group I intron in the mitochondrial cox1 gene during evolution of the Araceae family. Mol. Biol. Evol. 16, 1155-1165.

Clapp, J.P., Young, J.P.W., Merryweather, J.W., Fitter, A.H., 1995. Diversity of fungal symbionts in arbuscular mycorrhizas from a natural community. New Phytol. 130, 259-265.

Clark, R.B., 1997. Arbuscular mycorrhizal adaptation, spore germination, root colonization, and host plant growth and mineral acquisition at low pH. Plant and Soil 192, 15-22.

Colwell, R.K. 2005. EstimateS: Statistical estimation of species richness and shared species from samples.

Corkidi, L., Allen, E.B., Merhaut, D., Allen, M.F., Downer, J., Bohn, J., Evans, M., 2004. Assessing the infectivity of commercial mycorrhizal inoculants in plant nursery conditions. J. Envir. Hort. 22, 149-154.

Croll, D., Corradi, N., Gamper, H.A., Sanders, I.R., 2008a. Multilocus genotyping of arbuscular mycorrhizal fungi and marker suitability for population genetics. New Phytol. 180, 564-568.

Croll, D., Wille, L., Gamper, H.A., Mathimaran, N., Lammers, P.J., Corradi, N., Sanders, I.R., 2008b. Genetic diversity and host plant preferences revealed by simple sequence repeat and mitochondrial markers in a population of the arbuscular mycorrhizal fungus *Glomus intraradices*. New Phytol. 178, 672-687.

Croll, D., Giovannetti, M., Koch, A.M., Sbrana, C., Ehinger, M., Lammers, P.J., Sanders, I.R., 2009. Nonself vegetative fusion and genetic exchange in the arbuscular mycorrhizal fungus *Glomus intraradices*. New Phytol. 181, 924-937.

Croll, D., Sanders, I.R., 2009. Recombination in *Glomus intraradices*, a supposed ancient asexual arbuscular mycorrhizal fungus. BMC Evol. Biol. 9,

Cusimano, N., Zhang, L.B., Renner, S.S., 2008. Reevaluation of the cox1 group I intron in Araceae and angiosperms indicates a history dominated by loss rather than horizontal transfer. Mol. Biol. Evol. 25, 265-276.

Dalgaard, J.Z., Klar, A.J., Moser, M.J., Holley, W.R., Chatterjee, A., Mian, I.S., 1997. Statistical modeling and analysis of the LAGLIDADG family of site-specific endonucleases and identification of an intein that encodes a site-specific endonuclease of the HNH family. Nucleic Acids Res. 25, 4626-4638.

de Souza, F.A., Kowalchuk, G.A., Leeflang, P., van Veen, J.A., Smit, E., 2004. PCR-denaturing gradient gel electrophoresis profiling of inter- and intraspecies 18S rRNA gene sequence heterogeneity is an accurate and sensitive method to assess species diversity of arbuscular mycorrhizal fungi of the genus *Gigaspora*. Appl. Environ. Microbiol. 70, 1413-1424.

den Bakker, H.C., VanKuren, N.W., Morton, J.B., Pawlowska, T.E., 2010. Clonality and recombination in the life history of an asexual arbuscular mycorrhizal fungus. Mol. Biol. Evol. doi:10.1093/molbev/msq155

Dickson, S., Smith, F.A., Smith, S.E., 2007. Structural differences in arbuscular mycorrhizal symbioses: more than 100 years after Gallaud, where next? Mycorrhiza 17, 375-393.

Doner, L.W., Bécard, G., 1991. Solubilization of gellan gels by chelation of cations. Biotechnol. Tech. 5, 25-28.

Esch, H., Hundeshagen, B., Schneiderpoetsch, H., Bothe, H., 1994. Demonstration of abscisic-acid in spores and hyphae of the arbuscular-mycorrhizal fungus *Glomus* and in the N-2-Fixing Cyanobacterium *Anabaena variabilis*. Plant Sci. 99, 9-16.

Fitter, A.H., Gilligan, C.A., Hollingworth, K., Kleczkowski, A., Twyman, R.M., Pitchford, J.W., 2005. Biodiversity and ecosystem function in soil. Funct. Ecol. 19, 369-377.

Formey, D., Roux, C. 2009. Organization of the respiratory chain of *Glomus intraradices* deduced from genome analysis. In *"Beyond The Roots" 6th International Conference On Mycorrhiza (ICOM6)*. Belo Horizonte, Brazil

Frézal, L., Leblois, R., 2008. Four years of DNA barcoding: Current advances and prospects. Infect. Genet. Evol. 8, 727-736.

Gamalero, E., Trotta, A., Massa, N., Copetta, A., Martinotti, M.G., Berta, G., 2004. Impact of two fluorescent pseudomonads and an arbuscular mycorrhizal fungus on tomato plant growth, root architecture and P acquisition. Mycorrhiza 14, 185-192.

Gamper, H.A., van der Heijden, M.G.A., Kowalchuk, G.A., 2010. Molecular trait indicators: moving beyond phylogeny in arbuscular mycorrhizal ecology. New Phytol. 185, 67-82.

Gandolfi, A., Sanders, I.R., Rossi, V., Menozzi, P., 2003. Evidence of recombination in putative ancient asexuals. Mol. Biol. Evol. 20, 754-761.

Geml, J., Laursen, G.A., O'Neill, K., Nusbaum, H.C., Taylor, D.L., 2006. Beringian origins and cryptic speciation events in the fly agaric (*Amanita muscaria*). Mol. Ecol. 15, 225-239.

Genre, A., Chabaud, M., Timmers, T., Bonfante, P., Barker, D.G., 2005. Arbuscular mycorrhizal fungi elicit a novel intracellular apparatus in *Medicago truncatula* root epidermal cells before infection. Plant Cell 17, 3489-3499.

Gerdemann, J.W., Trappe, J.M., 1974. Endogonaceae in the Pacific Northwest. Mycologia Mem. 5, 1-76.

Germino, M.J., Wraith, J.M., 2003. Plant water relations influence carbon gain in a grass occurring along sharp gradients of soil temperature. New Phytol. 157, 241-250.

Gimble, F.S., 2000. Invasion of a multitude of genetic niches by mobile endonuclease genes. FEMS Microbiol. Lett. 185, 99-107.

Giovannetti, M., Azzolini, D., Citernesi, A.S., 1999. Anastomosis formation and nuclear and protoplasmic exchange in arbuscular mycorrhizal fungi. Appl. Environ. Microbiol. 65, 5571-5575.

Giovannetti, M., Sbrana, C., Strani, P., Agnolucci, M., Rinaudo, V., Avio, L., 2003. Genetic diversity of isolates of *Glomus mosseae* from different geographic areas detected by vegetative compatibility testing and biochemical and molecular analysis. Appl. Environ. Microbiol. 69, 616-624.

Giovannetti, M., Sbrana, C., Avio, L., Strani, P., 2004. Patterns of below-ground plant interconnections established by means of arbuscular mycorrhizal networks. New Phytol. 164, 175-181.

Goddard, M.R., Burt, A., 1999. Recurrent invasion and extinction of a selfish gene. Proc. Natl. Acad. Sci. USA 96, 13880-13885.

Gonzalez, P., Barroso, G., Labarere, J., 1999. Molecular gene organisation and secondary structure of the mitochondrial large subunit ribosomal RNA from the cultivated Basidiomycota *Agrocybe aegerita*: a 13 kb gene possessing six unusual nucleotide extensions and eight introns. Nucleic Acids Res. 27, 1754-1761.

Goudet, J., 2005. HIERFSTAT, a package for R to compute and test hierarchical F-statistics. Mol. Ecol. Notes 5, 184-186.

Graham, J.H., Hodge, N.C., Morton, J.B., 1995. Fatty acid methyl ester profiles for characterization of glomalean fungi and their endomycorrhizae. Appl. Environ. Microbiol. 61, 58-64.

Hall, T.A. 1999. Bioedit: user-friendly biological sequence alignment and analysis program for Windows 95/98/NT. Raleigh, NC, USA: North Carolina State University.

Hancock, J.M., Dover, G.A., 1988. Molecular Coevolution among Cryptically Simple Expansion Segments of Eukaryotic 26S/28S Ribosomal-RNAs. Mol. Biol. Evol. 5, 377-391.

Haugen, P., Bhattacharya, D., 2004. The spread of LAGLIDADG homing endonuclease genes in rDNA. Nucleic Acids Res. 32, 2049-2057.

Haugen, P., Reeb, V., Lutzoni, F., Bhattacharya, D., 2004. The evolution of homing endonuclease genes and group I introns in nuclear rDNA. Mol. Biol. Evol. 21, 129-140.

Haugen, P., Bhattacharya, D., Palmer, J.D., Turner, S., Lewis, L.A., Pryer, K.M., 2007. Cyanobacterial ribosomal RNA genes with multiple, endonuclease-encoding group 1 introns. BMC Evol. Biol. 7, 9.

Hawksworth, D.L., 2001. The magnitude of fungal diversity: the 1.5 million species estimate revisited. Mycol. Res. 105, 1422-1432.

Heath, P.J., Stephens, K.M., Monnat, R.J., Stoddard, B.L., 1997. The structure of I-CreI, a Group I intron-encoded homing endonuclease. Nat. Struct. Biol. 4, 468-476.

Hebert, P.D.N., Cywinska, A., Ball, S.L., DeWaard, J.R., 2003. Biological identifications through DNA barcodes. Proc. R. Soc. Lond. B Biol. Sci. 270, 313-321.

Helber, N., Requena, N., 2008. Expression of the fluorescence markers DsRed and GFP fused to a nuclear localization signal in the arbuscular mycorrhizal fungus *Glomus intraradices*. New Phytol. 177, 537-548.

Helgason, T., Daniell, T.J., Husband, R., Fitter, A.H., Young, J.P.W., 1998a. Ploughing up the wood-wide web? Nature 394, 431-431.

Helgason, T., Daniell, T.J., Husband, R., Fitter, A.H., Young, J.P.W., 1998b. Ploughing up the wood-wide web? Nature 394, 431.

Helgason, T., Fitter, A.H., Young, J.P.W., 1999. Molecular diversity of arbuscular mycorrhizal fungi colonising *Hyacinthoides non-scripta* (bluebell) in a seminatural woodland. Mol. Ecol. 8, 659-666.

Helgason, T., Merryweather, J.W., Denison, J., Wilson, P., Young, J.P.W., Fitter, A.H., 2002. Selectivity and functional diversity in arbuscular mycorrhizas of co-occurring fungi and plants from a temperate deciduous woodland. J. Ecol. 90, 371-384.

Hempel, S., Renker, C., Buscot, F., 2007. Differences in the species composition of arbuscular mycorrhizal fungi in spore, root and soil communities in a grassland ecosystem. Environ. Microbiol. 9, 1930-1938.

Henley, R.W., Hedenquist, J.W., Roberts, P.J. 1986. Guide to the active epithermal (geothermal) systems and precious metal deposits of New Zealand. In *Monograph Series on Mineral Deposits*. Berlin: Gebrüder Borntraeger.

Hijri, I., Sýkorová, Z., Oehl, F., Ineichen, K., Mader, P., Wiemken, A., Redecker, D., 2006. Communities of arbuscular mycorrhizal fungi in arable soils are not necessarily low in diversity. Mol. Ecol. 15, 2277-2289.

Hijri, M., Sanders, I.R., 2005. Low gene copy number shows that arbuscular mycorrhizal fungi inherit genetically different nuclei. Nature 433, 160-163.

Hovmøller, M.S., Yahyaoui, A.H., Milus, E.A., Justesen, A.F., 2008. Rapid global spread of two aggressive strains of a wheat rust fungus. Mol. Ecol. 17, 3818-3826.

Husband, R., Herre, E.A., Turner, S.L., Gallery, R., Young, J.P.W., 2002. Molecular diversity of arbuscular mycorrhizal fungi and patterns of host association over time and space in a tropical forest. Mol. Ecol. 11, 2669-2678.

Huson, D.H., Bryant, D., 2006. Application of phylogenetic networks in evolutionary studies. Mol. Biol. Evol. 23, 254-267.

James, T.Y., Kauff, F., Schoch, C.L., Matheny, P.B., Hofstetter, V., Cox, C.J., Celio, G., Gueidan, C., Fraker, E., Miadlikowska, J., Lumbsch, H.T., Rauhut, A., Reeb, V., Arnold, A.E., Amtoft, A., Stajich, J.E., Hosaka, K., Sung, G.H., Johnson, D., O'Rourke, B., Crockett, M., Binder, M., Curtis, J.M., Slot, J.C., Wang, Z., Wilson, A.W., Schussler, A., Longcore, J.E., O'Donnell, K., Mozley-Standridge, S., Porter, D., Letcher, P.M., Powell, M.J., Taylor, J.W., White, M.M., Griffith, G.W., Davies, D.R., Humber, R.A., Morton, J.B., Sugiyama, J., Rossman, A.Y., Rogers, J.D., Pfister, D.H., Hewitt, D., Hansen, K., Hambleton, S., Shoemaker, R.A., Kohlmeyer, J., Volkmann-Kohlmeyer, B., Spotts, R.A., Serdani, M., Crous, P.W., Hughes, K.W., Matsuura, K., Langer, E., Langer, G., Untereiner, W.A., Lucking, R., Budel, B., Geiser, D.M., Aptroot, A., Diederich, P., Schmitt, I., Schultz, M., Yahr, R., Hibbett, D.S., Lutzoni,

F., McLaughlin, D.J., Spatafora, J.W., Vilgalys, R., 2006. Reconstructing the early evolution of Fungi using a six-gene phylogeny. Nature 443, 818-822.

Jansa, J., Mozafar, A., Anken, T., Ruh, R., Sanders, I.R., Frossard, E., 2002a. Diversity and structure of AMF communities as affected by tillage in a temperate soil. Mycorrhiza 12, 225-234.

Jansa, J., Mozafar, A., Banke, S., McDonald, B.A., Frossard, E., 2002b. Intra- and intersporal diversity of ITS rDNA sequences in *Glomus intraradices* assessed by cloning and sequencing, and by SSCP analysis. Mycol. Res. 106, 670-681.

Jansa, J., Mozafar, A., Kuhn, G., Anken, T., Ruh, R., Sanders, I.R., Frossard, E., 2003. Soil tillage affects the community structure of mycorrhizal fungi in maize roots. Ecol. Appl. 13, 1164-1176.

Johnson, D., Vandenkoornhuyse, P.J., Leake, J.R., Gilbert, L., Booth, R.E., Grime, J.P., Young, J.P.W., Read, D.J., 2004. Plant communities affect arbuscular mycorrhizal fungal diversity and community composition in grassland microcosms. New Phytol. 161, 503-515.

Jurica, M.S., Stoddard, B.L., 1999. Homing endonucleases: structure, function and evolution. Cell. Mol. Life Sci. 55, 1304-1326.

Keese, P.K., Gibbs, A., 1992. Origins of genes - Big-bang or continuous creation. Proc. Natl. Acad. Sci. USA 89, 9489-9493.

Kjøller, R., Rosendahl, S., 2000. Detection of arbuscular mycorrhizal fungi (Glomales) in roots by nested PCR and SSCP (single stranded conformation polymorphism). Plant Soil 226, 189-196.

Klironomos, J.N., McCune, J., Hart, M., Neville, J., 2000. The influence of arbuscular mycorrhizae on the relationship between plant diversity and productivity. Ecol. Lett. 3, 137-141.

Klironomos, J.N., 2003. Variation in plant response to native and exotic arbuscular mycorrhizal fungi. Ecology 84, 2292-2301.

Koch, A.M., Kuhn, G., Fontanillas, P., Fumagalli, L., Goudet, J., Sanders, I.R., 2004. High genetic variability and low local diversity in a population of arbuscular mycorrhizal fungi. Proc. Natl. Acad. Sci. USA 101, 2369-2374.

Koch, A.M., Croll, D., Sanders, I.R., 2006. Genetic variability in a population of arbuscular mycorrhizal fungi causes variation in plant growth. Ecol. Lett. 9, 103-110.

Korber, B. 2000. In: AG Rodrigo, GH Learn, Computational analysis of HIV molecular sequences. Kluwer Academic Publishers, Dordrecht, Netherlands, pp. 55-72.

Krebs, C.J. 1994. *Ecology*. New York: HarperCollins College

Kretzer, A.M., Dunham, S., Molina, R., Spatafora, J.W., 2005. Patterns of vegetative growth and gene flow in *Rhizopogon vinicolor* and *R. vesiculosus* (Boletales, Basidiomycota). Mol. Ecol. 14, 2259-2268.

Krüger, M., Stockinger, H., Krüger, C., Schüßler, A., 2009. DNA-based species level detection of Glomeromycota: one PCR primer set for all arbuscular mycorrhizal fungi. New Phytol. 183, 212-223.

Kuhn, G., Hijri, M., Sanders, I.R., 2001. Evidence for the evolution of multiple genomes in arbuscular mycorrhizal fungi. Nature 414, 745-748.

Lanfranco, L., Delpero, M., Bonfante, P., 1999. Intrasporal variability of ribosomal sequences in the endomycorrhizal fungus *Gigaspora margarita*. Mol. Ecol. 8, 37-45.

Lang, B.F., Laforest, M.J., Burger, G., 2007. Mitochondrial introns: a critical view. Trends Genet. 23, 119-125.

Lang, B.F., Hijri, M., 2009. The complete *Glomus intraradices* mitochondrial genome sequence - a milestone in mycorrhizal research. New Phytol. 183, 4-6.

Lee, J., Young, J.P.W., 2009. The mitochondrial genome sequence of the arbuscular mycorrhizal fungus *Glomus intraradices* isolate 494 and implications for the phylogenetic placement of *Glomus*. New Phytol. 183, 200-211.

Lekberg, Y., Koide, R.T., Rohr, J.R., Aldrich-Wolfe, L., Morton, J.B., 2007. Role of niche restrictions and dispersal in the composition of arbuscular mycorrhizal fungal communities. J. Ecol. 95, 95-105.

Liu, Y., Steenkamp, E.T., Brinkmann, H., Forget, L., Philippe, H., Lang, B.F., 2009. Phylogenomic analyses predict sistergroup relationship of nucleariids and Fungi and paraphyly of zygomycetes with significant support. BMC Evol. Biol. 9:272.

Lloyd MacGilp, S.A., Chambers, S.M., Dodd, J.C., Fitter, A.H., Walker, C., Young, J.P.W., 1996. Diversity of the ribosomal internal transcribed spacers within and among isolates of *Glomus mosseae* and related mycorrhizal fungi. New Phytol. 133, 103-111.

Lucas, P., Otis, C., Mercier, J.P., Turmel, M., Lemieux, C., 2001. Rapid evolution of the DNA-binding site in LAGLIDADG homing endonucleases. Nucleic Acids Res. 29, 960-969.

Marchler-Bauer, A., Anderson, J.B., Chitsaz, F., Derbyshire, M.K., DeWeese-Scott, C., Fong, J.H., Geer, L.Y., Geer, R.C., Gonzales, N.R., Gwadz, M., He, S., Hurwitz, D.I., Jackson, J.D., Ke, Z., Lanczycki, C.J., Liebert, C.A., Liu, C., Lu, F., Lu, S., Marchler, G.H., Mullokandov, M., Song, J.S., Tasneem, A., Thanki, N., Yamashita, R.A., Zhang, D., Zhang, N., Bryant, S.H., 2009. CDD: specific functional annotation with the Conserved Domain Database. Nucleic Acids Res. 37, D205-D210.

Marinoni, G., Manuel, M., Petersen, R.F., Hvidtfeldt, J., Sulo, P., Piskur, J., 1999. Horizontal transfer of genetic material among *Saccharomyces* yeasts. J. Bacteriol. 181, 6488-6496.

Marquez, L.M., Redman, R.S., Rodriguez, R.J., Roossinck, M.J., 2007. A virus in a fungus in a plant: Three-way symbiosis required for thermal tolerance. Science 315, 513-515.

Martin, F., Tuskan, G.A., DiFazio, S.P., Lammers, P., Newcombe, G., Podila, G.K., 2004. Symbiotic sequencing for the *Populus* mesocosm. New Phytol. 161, 330-335.

Martin, F., Gianinazzi-Pearson, V., Hijri, M., Lammers, P., Requena, N., Sanders, I.R., Shachar-Hill, Y., Shapiro, H., Tuskan, G.A., Young, J.P.W., 2008. The long hard road to a completed *Glomus intraradices* genome. New Phytol. 180, 747-750.

Mathimaran, N., Falquet, L., Ineichen, K., Picard, C., Redecker, D., Boller, T., Wiemken, A., 2008a. Microsatellites for disentangling underground networks: Strain-specific identification of *Glomus intraradices*, an arbuscular mycorrhizal fungus. Fungal Genet. Biol. 45, 812-817.

Mathimaran, N., Falquet, L., Ineichen, K., Picard, C., Redecker, D., Wiemken, A., Boller, T., 2008b. Unexpected vagaries of microsatellite loci in *Glomus intraradices*: length polymorphisms are rarely caused by variation in repeat number only. New Phytol. 180, 568-570.

Morton, J.B., 1988. Taxonomy of VA mycorrhizal fungi: Classification, nomenclature, and identification. Mycotaxon 32, 267-324.

Morton, J.B., Bever, J.D., Pfleger, F.L., 1997. Taxonomy of *Acaulospora gerdemannii* and *Glomus leptotichum*, synanamorphs of an arbuscular mycorrhizal fungus in Glomales. Mycol. Res. 101, 625-631.

Munkvold, L., Kjoller, R., Vestberg, M., Rosendahl, S., Jakobsen, I., 2004. High functional diversity within species of arbuscular mycorrhizal fungi. New Phytol. 164, 357-364.

Navajas, M., Navia, D., 2010. DNA-based methods for eriophyoid mite studies: review, critical aspects, prospects and challenges. Exp. Appl. Acarol. 51, 257-271.

Neuveglise, C., Brygoo, Y., Riba, G., 1997. 28s rDNA group-I introns: A powerful tool for identifying strains of *Beauveria brongniartii*. Mol. Ecol. 6, 373-381.

Ochsner, A. 2002. *Analysen der mitochondriellen grossen Untereinheit der ribosomalen DNA von Pilzen der arbuskulären Mykorrhiza*. MSc thesis. University of Basel, Basel.

Oehl, F., Sieverding, E., Ineichen, K., Mader, P., Boller, T., Wiemken, A., 2003. Impact of land use intensity on the species diversity of arbuscular mycorrhizal fungi in agroecosystems of Central Europe. Appl. Environ. Microbiol. 69, 2816-2824.

Öpik, M., Moora, M., Liira, J., Koljalg, U., Zobel, M., Sen, R., 2003. Divergent arbuscular mycorrhizal fungal communities colonize roots of *Pulsatilla* spp. in boreal Scots pine forest and grassland soils. New Phytol. 160, 581-593.

Öpik, M., Moora, M., Liira, J., Zobel, M., 2006. Composition of root-colonizing arbuscular mycorrhizal fungal communities in different ecosystems around the globe. J. Ecol. 94, 778-790.

Öpik, M., Vanatoa, A., Vanatoa, E., Moora, M., Davison, J., Kalwij, J.M., Reier, Ü., Zobel, M., 2010. The online database MaarjAM reveals global and ecosystemic distribution patterns in arbuscular mycorrhizal fungi (Glomeromycota). New Phytol.

Parniske, M., 2004. Molecular genetics of the arbuscular mycorrhizal symbiosis. Curr. Opin. Plant Biol. 7, 414-421.

Pawlowska, T.E., Taylor, J.W., 2004. Organization of genetic variation in individuals of arbuscular mycorrhizal fungi. Nature 427, 733-737.

Piotrowski, J.S., Denich, T., Klironomos, J.N., Graham, J.M., Rillig, M.C., 2004. The effects of arbuscular mycorrhizas on soil aggregation depend on the interaction between plant and fungal species. New Phytol. 164, 365-373.

Pivato, B., Offre, P., Marchelli, S., Barbonaglia, B., Mougel, C., Lemanceau, P., Berta, G., 2009. Bacterial effects on arbuscular mycorrhizal fungi and mycorrhiza development as influenced by the bacteria, fungi, and host plant. Mycorrhiza 19, 81-90.

Pleines, T., Jakob, S.S., Blattner, F.R., 2009. Application of non-coding DNA regions in intraspecific analyses. Plant Syst. Evol. 282, 281-294.

Posada, D. 2004. Modeltest 3.5. Vigo, Spain: Facultad de Biologia, Universidad de Vigo.

Qiu, Y.L., Cho, Y., Cox, J.C., Palmer, J.D., 1998. The gain of three mitochondrial introns identifies liverworts as the earliest land plants. Nature 394, 671-674.

Raab, P.A., Brennwald, A., Redecker, D., 2005. Mitochondrial large ribosomal subunit sequences are homogeneous within isolates of *Glomus* (arbuscular mycorrhizal fungi, Glomeromycota). Mycol. Res. 109, 1315-1322.

Raab, P.A. 2007. *Development of new molecular markers for phylogeny and molecular identification of arbuscular mycorrhizal fungi (Glomeromycota)*. PhD thesis. University of Basel, Basel.

Redecker, D., Thierfelder, H., Walker, C., Werner, D., 1997. Restriction analysis of PCR-amplified internal transcribed spacers of ribosomal DNA as a tool for species identification in different genera of the order glomales. Appl. Environ. Microbiol. 63, 1756-1761.

Redecker, D., 2000. Specific PCR primers to identify arbuscular mycorrhizal fungi within colonized roots. Mycorrhiza 10, 73-80.

Redecker, D., Kodner, R., Graham, L.E., 2000. Glomalean fungi from the Ordovician. Science 289, 1920-1921.

Redecker, D. 2006. In: JR J Cooper, Molecular techniques for soil and rhizosphere microorganisms. Oxfordshire, UK, pp. 198-212.

Redecker, D., Raab, P., 2006. Phylogeny of the Glomeromycota (arbuscular mycorrhizal fungi): recent developments and new gene markers. Mycologia 98, 885-895.

Reeb, V., Haugen, P., Bhattacharya, D., Lutzoni, F., 2007. Evolution of *Pleopsidium* (Lichenized ascomycota) S943 group I introns and the phylogeography of an intron-encoded putative homing endonuclease. J. Mol. Evol. 64, 285-298.

Renker, C., Blanke, V., Buscot, F., 2005. Diversity of arbuscular mycorrhizal fungi in grassland spontaneously developed on area polluted by a fertilizer plant. Environ. Pollut. 135, 255-266.

Repetto, O., Massa, N., Gianinazzi-Pearson, V., Dumas-Gaudot, E., Berta, G., 2007. Cadmium effects on populations of root nuclei in two pea genotypes inoculated or not with the arbuscular mycorrhizal fungus *Glomus mosseae*. Mycorrhiza 17, 111-120.

Rodman, A., Shovic, H., Thoma, D. 1996. *Soils of Yellowstone National Park.* Yellowstone National Park: Yellowstone Center for Resources YCR-NRSR-96-2, National Park Service.

Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572-1574.

Rose, A.B., 2008. Intron-Mediated Regulation of Gene Expression. Curr. Top. Microbiol. 326, 277-290.

Rosenberg, M.S. 2001. PASSaGE. Pattern Analysis, Spatial Statistics, and Geographic Exegesis. Department of Biology, Arizona State University, Temper.

Rosendahl, S., Stukenbrock, E.H., 2004. Community structure of arbuscular mycorrhizal fungi in undisturbed vegetation revealed by analyses of LSU rDNA sequences. Mol. Ecol. 13, 3179-3186.

Rosendahl, S., 2008. Communities, populations and individuals of arbuscular mycorrhizal fungi. New Phytol. 178, 253-266.

Rosendahl, S., Matzen, H.B., 2008. Genetic structure of arbuscular mycorrhizal populations in fallow and cultivated soils. New Phytol. 179, 1154-1161.

Rosendahl, S., McGee, P., Morton, J.B., 2009. Lack of global population genetic differentiation in the arbuscular mycorrhizal fungus *Glomus mosseae* suggests a recent range expansion which may have coincided with the spread of agriculture. Mol. Ecol. 18, 4316-4329.

Sanchez-Puerta, M.V., Cho, Y., Mower, J.P., Alverson, A.J., Palmer, J.D., 2008. Frequent, phylogenetically local horizontal transfer of the cox1 group I intron in flowering plant mitochondria. Mol. Biol. Evol. 25, 1762-1777.

Sandegren, L., Sjoberg, B.M., 2004. Distribution, sequence homology, and homing of group I Introns among T-even-like bacteriophages - Evidence for recent transfer of old introns. J. Biol. Chem. 279, 22218-22227.

Sanders, I.R., Alt, M., Groppe, K., Boller, T., Wiemken, A., 1995. Identification of ribosomal DNA polymorphisms among and within spores of the Glomales: Application to studies on the genetic diversity of arbuscular mycorrhizal fungal communities. New Phytol. 130, 419-427.

Saville, B.J., Kohli, Y., Anderson, J.B., 1998. mtDNA recombination in a natural population. Proc. Natl. Acad. Sci. USA 95, 1331-1335.

Schenck, N.C., Smith, G.S., 1982. Additional new and unreported species of mycorrhizal fungi (Endogonaceae) from Florida. Mycologia 74, 77-92.

Schenck, N.C., Perez, Y. 1990. *Manual for the identification of VA mycorrhizal fungi*. Gainesville: Synergistic Publications.

Scheublin, T.R., Ridgway, K.P., Young, J.P.W., van der Heijden, M.G.A., 2004. Nonlegumes, legumes, and root nodules harbor different arbuscular mycorrhizal fungal communities. Appl. Environ. Microbiol. 70, 6240-6246.

Schreiner, R.P., Bethlenfalvay, G.J., 1995. Mycorrhizal interactions in sustainable agriculture. Crit. Rev. Biotechnol. 15, 271-285.

Schüßler, A., Gehrig, H., Schwarzott, D., Walker, C., 2001a. Analysis of partial Glomales SSU rRNA gene sequences: implications for primer design and phylogeny. Mycol. Res. 105, 5-15.

Schüßler, A., Schwarzott, D., Walker, C., 2001b. A new fungal phylum, the Glomeromycota: phylogeny and evolution. Mycol. Res. 105, 1414-1421.

Schwarzott, D., Walker, C., Schüßler, A., 2001. *Glomus*, the largest genus of the arbuscular mycorrhizal fungi (Glomales), is non-monophyletic. Mol. Phylogenet. Evol. 21, 190-197.

Sears, B.B., 1980. Changes in chloroplast genome composition and recombination during the maturation of zygospores of *Chlamydomonas reinhardtii*. Curr. Genet. 2, 1-8.

Seifert, K.A., Samson, R.A., Dewaard, J.R., Houbraken, J., Levesque, C.A., Moncalvo, J.M., Louis-Seize, G., Hebert, P.D.N., 2007. Prospects for fungus identification using C01 DNA barcodes, with *Penicillium* as a test case. Proc. Natl. Acad. Sci. USA 104, 3901-3906.

Selosse, M.A., Le Tacon, F., 1998. The land flora: a phototroph-fungus partnership? Trends Ecol. Evol. 13, 15-20.

Shibata, T., Ling, F., 2007. DNA recombination protein-dependent mechanism of homoplasmy and its proposed functions. Mitochondrion 7, 17-23.

Silva, G.H., Belfort, M., Wende, W., Pingoud, A., 2006. From monomeric to homodimeric endonucleases and back: Engineering novel specificity of LAGLIDADG enzymes. J. Mol. Biol. 361, 744-754.

Smith, S.E., Read, D.J. 1997. *Mycorrhizal symbiosis*. London: Academic Press.

Smith, S.E., Smith, F.A., Jakobsen, I., 2004. Functional diversity in arbuscular mycorrhizal (AM) symbioses: the contribution of the mycorrhizal P uptake pathway is not correlated with mycorrhizal responses in growth or total P uptake. New Phytol. 162, 511-524.

Smith, S.E., Read, D.J. 2008. *Mycorrhizal symbiosis*. London: Academic Press

Sor, F., Fukuhara, H., 1983. Complete DNA-sequence coding for the large ribosomal RNA of yeast mitochondria. Nucleic Acids Res. 11, 339-348.

Stockinger, H., Walker, C., Schüßler, A., 2009. '*Glomus intraradices* DAOM197198', a model fungus in arbuscular mycorrhiza research, is not *Glomus intraradices*. New Phytol. 183, 1176-1187.

Stockinger, H., Kruger, M., Schüßler, A., 2010. DNA barcoding of arbuscular mycorrhizal fungi. New Phytol. 187, 461-474.

Stout, R.G., Kerstetter, T., Summers, M., 1997. Plants in extreme environments: heat- and acid-tolerant vascular plants and mosses found in Yellowstone's geothermal environments. Plant Physiol. 114, 498-498.

Streitwolf-Engel, R., Boller, T., Wiemken, A., Sanders, I.R., 1997. Clonal growth traits of two *Prunella* species are determined by co-occurring arbuscular mycorrhizal fungi from a calcareous grassland. J. Ecol. 85, 181-191.

Stukenbrock, E.H., Rosendahl, S., 2005a. Clonal diversity and population genetic structure of arbuscular mycorrhizal fungi (*Glomus* spp.) studied by multilocus genotyping of single spores. Mol. Ecol. 14, 743-752.

Stukenbrock, E.H., Rosendahl, S., 2005b. Development and amplification of multiple co-dominant genetic markers from single spores of arbuscular mycorrhizal fungi by nested multiplex PCR. Fungal Genet. Biol. 42, 73-80.

Stürmer, S.L., Morton, J.B., 1999. Taxonomic reinterpretation of morphological characters in Acaulosporaceae based on developmental patterns. Mycologia 91, 849-857.

Swofford, D.L. 2001. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Sunderland, MA, USA: Sinauer Associates.

Sýkorová, Z., Ineichen, K., Wiemken, A., Redecker, D., 2007a. The cultivation bias: different communities of arbuscular mycorrhizal fungi detected in roots from the field, from bait plants transplanted to the field, and from a greenhouse trap experiment. Mycorrhiza 18, 1-14.

Sýkorová, Z., Wiemken, A., Redecker, D., 2007b. Cooccurring *Gentiana verna* and *Gentiana acaulis* and their neighboring plants in two swiss upper montane meadows harbor distinct arbuscular mycorrhizal fungal communities. Appl. Environ. Microbiol. 73, 5426-5434.

Tehler, A., Little, D.P., Farris, J.S., 2003. The full-length phylogenetic tree from 1551 ribosomal sequences of chitinous fungi, Fungi. Mycol. Res. 107, 901-916.

ter Braak, C.F.J., Smilauer, P. 2004. CANOCO Reference Manual and CanoDraw for Windows User's guide: Software for Canonical Community Ordination pp. Biometris, Wageningen, Netherlands. Ceske Budejovice, Czech Republic.

Tercek, M.T., Whitbeck, J.L., 2004. Heat avoidance life history strategy controls the distribution of geothermal *Agrostis* in Yellowstone. Ecology 85, 1955-1966.

Thiéry, O., Börstler, B., Ineichen, K., Redecker, D., 2010. Evolutionary dynamics of introns and homing endonuclease ORFs in a region of the large subunit of the mitochondrial rRNA in *Glomus* species (arbuscular mycorrhizal fungi, Glomeromycota). Mol. Phylogenet. Evol. 55, 599-610.

van der Heijden, M.G.A., Klironomos, J.N., Ursic, M., Moutoglis, P., Streitwolf-Engel, R., Boller, T., Wiemken, A., Sanders, I.R., 1998. Mycorrhizal fungal diversity determines plant biodiversity, ecosystem variability and productivity. Nature 396, 69-72.

van der Heijden, M.G.A., Scheublin, T.R., Brader, A., 2004. Taxonomic and functional diversity in arbuscular mycorrhizal fungi - is there any relationship? New Phytol. 164, 201-204.

van Tuinen, D., Jacquot, E., Zhao, B., Gollotte, A., Gianinazzi-Pearson, V., 1998. Characterization of root colonization profiles by a microcosm community odf arbuscular mycorrhizal fungi using 25 rDNA-targeted nested PCR. Mol. Ecol. 7, 879-887.

Vaughan, J.C., Mason, M.T., Sper-Whitis, G.L., Kuhlman, P., Palmer, J.D., 1995. Fungal origin by horizontal transfer of a plant mitochondrial group I intron in the chimeric coxI gene of *Peperomia*. J. Mol. Evol. 41, 563-572.

Vincze, T., Posfai, J., Roberts, R.J., 2003. NEBcutter: a program to cleave DNA with restriction enzymes. Nucleic Acids Res. 31, 3688-3691.

Voets, L., de la Providencia, I.E., Declerck, S., 2006. Glomeraceae and Gigasporaceae differ in their ability to form hyphal networks. New Phytol. 172, 185-188.

Waples, R.S., Gaggiotti, O., 2006. What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. Mol. Ecol. 15, 1419-1439.

Wolff, G., Burger, G., Lang, B.F., Kuck, U., 1993. Mitochondrial genes in the colorless alga *Prototheca wickerhamii* resemble plant genes in their exons but fungal genes in their introns. Nucleic Acids Res. 21, 719-726.

Woodson, S.A., Cech, T.R., 1989. Reverse self-splicing of the *Tetrahymena* group-I intron - Implication for the directionality of splicing and for intron transposition. Cell 57, 335-345.

Wu, B.Y., Hogetsu, T., Isobe, K., Ishii, R., 2007. Community structure of arbuscular mycorrhizal fungi in a primary successional volcanic desert on the southeast slope of Mount Fuji. Mycorrhiza 17, 495-506.

Wubet, T., Weiss, M., Kottke, I., Teketay, D., Oberwinkler, F., 2004. Molecular diversity of arbuscular mycorrhizal fungi in *Prunus africana*, an endangered medicinal tree species in dry Afromontane forests of Ethiopia. New Phytol. 161, 517-528.

Xu, J. 2005. *Evolutionary Genetics of Fungi*. Norfolk: Horizon Bioscience.

Yokoyama, T., Suzuki, T., 2008. Ribosomal RNAs are tolerant toward genetic insertions: evolutionary origin of the expansion segments. Nucleic Acids Res. 36, 3539-3551.

Yoshida, S., Parniske, M., 2005. Regulation of plant symbiosis receptor kinase through serine and threonine phosphorylation. J. Biol. Chem 280, 9203-9209.

Yoshida, S., Maruyama, S., Nozaki, H., Shirasu, K., 2010. Horizontal gene transfer by the parasitic plant *Striga hermonthica*. Science 328, 1128-1128.

Zischler, H., Geisert, H., Von Haeseler, A., Paabo, S., 1995. A nuclear 'fossil' of the mitochondrial D-loop and the origin of modern humans. Nature 378, 489-492.

Curriculum vitae

Personal

Name	Odile Thiéry
Address	5 rue de Montreux, 68 300 St Louis, France
E-mail	Odile.Thiery@unibas.ch
Birth date	29.10.1982

Education and qualifications

2006-2010	**PhD position** at the University of Basel, Basel, Switzerland; supervised by Prof. Dirk Redecker
2005-2006	**Master of Science Degree level 2 in Biochemistry, Molecular and Cell Biology** at the University of Burgundy, Dijon, France Major: Molecular and Cell Biology Optional courses: Plant Biology, Methodology, English
2004-2005	**Master of Science Degree level 1 in Biochemistry** at the University of Burgundy, Dijon, France Major: Molecular and Cell Biology Optional courses: Pharmacology, Bioinformatics, English Research and study work: nanosensors for the studying of biological macromolecules
2001-2004	**Bachelor of Science in Biochemistry** at the University of Burgundy, Dijon, France
2000-2001	**TOEFL** (263/300) at the test center of Puyallup (USA) Intensive English Program grade A^+ and Biology courses (5 credits) grade 3.4 at the University of Washington, Seattle, USA Autumn 2000 and Spring 2001: Outstanding Student

Training periods

2008 (1 week) — University of Tübingen, Tübingen, Germany
Project leader: Dr. Markus Göker

- Theme: Molecular phylogenetic reconstruction

2005-2006 (9 months) — National Institute of Research on Agronomy, Dijon, France
Project leader: Dr. Eliane Dumas-Gaudot

- Theme: The time course analysis of the accumulation of transcripts and/or proteins of two genes encoding respectively a nodulin-like and a blue copper protein during mycorrhizal symbiosis

2004 (2 months) — Unit of Lipids and Nutrition at the University of Burgundy, Dijon, France
Project leader: Dr. Françoise Leborgne

- Theme: Study of the influence of Ghrelin on the transcriptional level of Carnitin Palmitoyl Transferase I

International conferences

Thiéry, O., Börstler, B., Redecker, D. (2009). Mitochondrial large ribosomal subunit gene evolution in the *Glomus* group A lineage of arbuscular mycorrhizal fungi (AMF). Oral presentation at the "6th International Conference On Mycorrhiza (ICOM6), *Beyond the roots*", Belo Horizonte, Brazil.
Thiéry, O., Börstler, B., Redecker, D. (2008). Development of mitochondrial genes as molecular markers in the Glomeromycota. Poster presented at the **"**Plant-Microbial Interactions (PMI) 2008" in Kraków, Poland.
Thiéry, O., Redecker, D. (2007). Development of mitochondrial genes as molecular markers in the Glomeromycota. Poster presented at the "Annual Meeting of the Mycological Society of America (MSA) 2007" in Baton Rouge, Louisiana, USA.

Publications

Thiéry, O., Börstler, B., Ineichen, K., Redecker, D. (2010). Evolutionary dynamics of introns and homing endonuclease ORFs in a region of the large subunit of the mitochondrial rRNA in *Glomus* species (arbuscular mycorrhizal fungi, Glomeromycota). Molecular Phylogenetics and Evolution 55: 599-610.
Börstler, B., **Thiéry, O.**, Sýkorová, Z., Berner, A., Redecker, D. (2010). Diversity of mitochondrial large subunit rDNA haplotypes of *Glomus intraradices* in two agricultural field experiments and two semi-natural grasslands. Molecular Ecology 19: 1497–1511.
Börstler, B., Raab, P., **Thiéry, O.**, Morton, J., Redecker, D. (2008). Genetic diversity of the arbuscular mycorrhizal fungus *Glomus intraradices* as determined by mitochondrial large subunit rRNA gene sequences is considerably higher than previously expected. New Phytologist 180: 452-465.
Nagaraj, S., Lei, Z., Watson, B., Sumner, L., Gallardo, K., Dumas-Gaudot, E., Recorbet, G., Robert, F., **Thiéry, O.**, Valot, B., Mathesius, U., Triplett, E. (2006). Elaboration of the chapter proteomics in the *Medicago truncatula* Handbook online.